



Volume 1, Nomor 1, Juni 2020, Hal 1-14

INFOTECH: Jurnal Informatika & Teknologi
p ISSN 2722-9378 | e ISSN 2722-9386

KOMPARASI AKURASI METODE K-NEAREST NEIGHBOR DAN C4.5 UNTUK CUSTOMER RELATIONSHIP MANAGEMENT PADA PERUSAHAAN PEMBIAYAAN

Mohamad Anas Sobarnas^{1*}, Iskandar²

^{1,2} Program Studi Teknik Informatika, Sekolah Tinggi Teknologi Muhammadiyah Cileungsi.

^{1,2} Jl. Angrek No.25, Perum. PTSC, Cileungsi, Bogor, Jawa Barat, Indonesia 16820

Informasi Artikel

Article History:

Submission: 25-05-2020

Revised: 15-06-2020

Accepted: 25-06-2020

Published: 30-06-2020

Kata Kunci:

CRM, C 4.5, Klasifikasi, K-NN,
Perusahaan Pembiayaan,

Keywords:

Teacher assistance, Mobile,
Geofencing, GPS, Wifi.

* Korespondensi:

Mohamad Anas Sobarnas
anas@sttmcileungsi.ac.id

Abstrak

Faktor utama pembiayaan bisnis adalah Layanan yang Mudah dan Cepat untuk pelanggan, Salah satu masalah perusahaan pembiayaan adalah kalah bersaing karena tidak adanya metode yang sistematis dalam menjaga hubungan yang baik dengan pelanggannya sehingga pelanggan tidak bertambah bahkan pelanggan lama pergi begitu saja. Salah satu metode dalam menjaga hubungan yang baik dengan pelanggan adalah dengan dibuatnya sistem komparasi untuk menguji tingkat kelayakan pelanggan, Sehingga pelanggan tetap loyal kepada Perusahaan. Adapun alat sebagai komparasi dalam aplikasi klasifikasi Pelanggan menggunakan algoritma Klasifikasi yaitu algoritma K-Nearest Neighbor (K-NN) dan C4.5. Setelah dilakukan pengujian perbandingan kedua algoritma diatas ditemukan akurasi algoritma C4.5 sebesar 93.10% dan nilai akurasi model untuk algoritma KKN sebesar 90.52% dengan selisih akurasi 2,58%. Pengujian aplikasi berbasis web menggunakan bahasa program ASP.Net C # membandingkan hasil akurasi dari perbandingan kedua algoritma diatas maka algoritma (C4.5) adalah menghasilkan nilai yang terbaik, kemudian dilakukan pengujian data sampling keakuratan aplikasi dengan jumlah pengujian data 122 sample didapatkan hasil 109.8 sample akurat atau sebesar 92% dalam menentukan loyalitas klasifikasi pelanggan. Sehingga dengan adanya sistem aplikasi komparasi pelanggan berbasis web ini bisa menjadi alat bantu perusahaan dalam menguji loyalitas pelanggan pada perusahaan.

ACCURACY COMPARATION OF K-NEAREST NEIGHBOR AND C4.5 METHOD FOR CUSTOMER RELATIONSHIP MANAGEMENT IN FINANCING COMPANIES

Abstract

The main factor of business financing is Easy and Fast Service for customers, One of the problems of finance companies is that they cannot compete because there is no systematic method of maintaining good relationships with their customers so that customers do not increase even old customers leave. One method of maintaining good relationships with customers is by making a comparative system to test the level of customer loyalty, so that customers remain loyal to the Company. The tool as a comparison in the Customer classification application uses the Classification algorithm, the algorithm. K-Nearest Neighbor (K-NN) and C4.5. After testing the comparison of the two algorithms above, it was found that the accuracy of the C4.5 algorithm was 93.10% and the accuracy of the model for the KKN algorithm was 90.52% with an



Infotech: Jurnal Informatika & Teknologi is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License. p ISSN 2722-9378 | e ISSN 2722-9386

accuracy difference of 2.58%. Testing a web-based application using the ASP.Net C # program language comparing the results of the accuracy of the comparison of the two algorithms above, the algorithm (C4.5) is to produce the best value, then testing the accuracy of the application data sampling with the number of test data 122 samples obtained 109.8 samples accurate or as much as 92% in determining customer classification loyalty. So that the web-based customer comparison application system can be a company tool in testing customer loyalty in the company.

1. PENDAHULUAN

Perusahaan pembiayaan adalah sektor penting dalam perekonomian dalam Industri lembaga pembiayaan, merupakan badan usaha diluar Bank dan Lembaga Keuangan Bukan Bank yang khusus didirikan untuk melakukan kegiatan usaha seperti sewa guna usaha, anjak piutang, usaha kartu kredit dan atau pembiayaan konsumen. Skema bisnis perusahaan pembiayaan didasari oleh adanya *underlying asset* hal ini membuat industri pembiayaan lebih dekat dengan konsumennya.

Dalam menjaga kestabilan dalam persaingan pada bisnis pembiayaan ini perlunya adanya upaya yang lebih komprehensif oleh pihak manajemen dalam rangka strategi untuk menjual produk dalam pembiayaan.

Teknik pengambilan keputusan yang didukung sistem informasi yang memadai baik didukung infrastruktur dan *basic knowledge* tentang data pada kemauan pelanggan yang cukup bisa menjadikan sebagai panduan pada manajemen dalam mengambil kebijakan terkait strategi bisnis yang akan dijalankan atau planning perusahaan.

Penelitian akan dilakukan dengan cara membandingkan dua algoritma klasifikasi berupa C4.5 dan K-Nearest Neighbor dengan aplikasi untuk mencari algoritma dengan akurasi terbaik. Hasil dari penelitian ini bisa dijadikan rujukan bagi seluruh pihak sebagai bahan pertimbangan dalam sistem pengambilan keputusan untuk mendapatkan *Grading Customer* terbaik.

2. METODE

2.1 Data Mining

Basis data (*Database*) adalah sekumpulan data yang saling berelasi antara satu dengan lainnya yang dilakukan proses komputerisasi sehingga menjadi informasi yang bisa berguna [1]. Manajemen data merupakan bagian dari manajemen sumber daya informasi yang membantu perusahaan agar sumber daya informasi yang dimilikinya mencerminkan secara akurat sistem fisik yang diwakilinya [2]. *Data Warehouse* adalah koleksi data yang mempunyai sifat berorientasi subjek terintegrasi dan konsisten, time variant dan bersifat tetap dan tidak berubah [3].

Definisi diatas menggambarkan konsep awal pengolahan data (data mining) yang bisa digunakan oleh perusahaan untuk membantu proses pengolahan data.

Data mining adalah analisis secara terstruktur dari kumpulan data yang banyak dan kompleks bertujuan untuk menemukan karakter pola dari hasil pengolahan datanya [4].

Data mining adalah kegiatan yang meliputi pengumpulan, pemakaian data historis untuk menemukan keteraturan, pola atau hubungan dalam set data berukuran besar. Keluaran data mining ini bisa dipakai untuk pengambilan keputusan masa yang akan datang [5].

Pengertian data mining dan KDD yaitu teknik data mining dalam pengolahan data dalam jumlah besar dengan tujuan menggali informasi yang terkandung dari basis data tersebut [6].

Adapun metode penelitian yang penulis gunakan menggunakan metode CRISP-DM melalui beberapa tahapan.



Gambar 1. Metode penelitian

2.2 Proses Data Mining

Secara sistematis, langkah-langkah utama yang terkandung di data mining terdiri [7].

a) Eksplorasi data awal.

Eksplorasi/pemrosesan awal data seperti *cleansing*, normalisasi, transformasi, data tidak valid.

b) Membangun model pola.

Tahapan selanjutnya adalah melakukan validasi dan analisis pada metode-metode yang digunakan seperti pengklasifikasian, pengklasteran, regresi analisis pada kluster pendeteksian anomali, analisis asosiasi, analisis pola sekuensial, dan sebagainya. Dalam beberapa referensi, deteksi anomali juga masuk dalam langkah eksplorasi. Akan tetapi, deteksi anomali juga dapat digunakan sebagai algoritma utama, terutama untuk mencari data-data yang spesial.

c) Penerapan.

Penerapan berarti menerapkan dari model pola baru dari karakteristik data yang ada sebagai hasil perkiraan masalah yang dihadapi.

2.3 Klasifikasi

Klasifikasi adalah sebuah proses analisa data yang menghasilkan model-model untuk menggambarkan kelas-kelas yang terkandung di dalam data. Model-model tersebut disebut *classifier*. Jadi, classifier inilah yang akan digunakan untuk menyusun kelas-kelas

yang terkandung didalam data. Ada banyak jenis algoritma klasifikasi, dua diantaranya adalah *Decision Tree* dan *K Nearest Neighbour (K-NN)*.

Klasifikasi didefinisikan sebagai bentuk analisis data untuk mengekstrak model yang akan digunakan untuk memprediksi label kelas [8].

2.4 Algoritma KNN

K-Nearest Neighbor (K-NN) termasuk kelompok *instance-based learning*. Dimana cara kerja algoritma ini adalah dengan cara mencari kelompok k objek pada data training yang mendekati dari pada kumpulan data [9].

Adapun tujuan dari algoritma klasifikasi KNN ini untuk mengklasifikasikan objek berdasarkan data training dan atribut, langkah-langkah algoritma KNN.

1. Tentukan parameter K
2. Tentukan jarak antara data yang akan dievaluasi dengan semua data training
3. Lalu urutkan hasil jarak secara menaik
4. Hitung jarak terdekat sampai dengan urutan ke-K
5. Lalu pasangkan kelas yang sesuai
6. Tentukan jumlah kelas dari tetangga yang paling dekat dan pastikan kelas tersebut adalah kelas dari data evaluasi.

Proses hitung dari rumus k-NN.

$$d_i = \sqrt{\sum_{i=1}^p (x_{2i} - x_{1i})^2} \quad (1)$$

Keterangan:

x_1 = Sampel Data x_2 = Data Pengujian i = Variabel Data

d = Jarak p = Dimensi Data

2.4.1 Algoritma C4.5

Algoritma C4.5 adalah teknik algoritma klasifikasi dalam membentuk model pohon keputusan. Sedang pohon keputusan dapat diartikan suatu cara untuk mengklarifikasi yang sangat kuat. Pohon keputusan diartikan bahwa kumpulan dari data kompleks dan besar menjadi part atau bagian terkecil dengan menerapkan *basic knowledge* keputusan [6].

Algoritma C4.5 adalah algoritma ekstensi dari ID3 yang lebih disukai karena memiliki kelebihan-kelebihan dalam pembentukan pola model keputusan dari karakter datanya dan juga memiliki performance yang lebih cepat dari algoritma yang sejenisnya.

Algoritma ini menggunakan rasio perolehan atau *Gain Ratio* dengan cara melakukan perhitungan nilai pada informasi dalam satuan bits dengan menggunakan metode entropi.

$$Entropy(S) = -p_{\oplus} \log_2 p_{\oplus} - p_{\ominus} \log_2 p_{\ominus} \quad (2)$$

S = Himpunan Kasus

P = Total Kasus dalam partisi ke- i

Setelah perhitungan entropi, dilakukan pula perhitungan Gain SA dimana, Gain SA ini adalah metode mencari informasi atribut A relatif pada data S, dimana rumus dari Gain SA.

$$Gain(S,A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} \quad (3)$$

Digunakan dalam proses klasifikasi atau clustering. Tujuan dari seleksi fitur ini adalah untuk mengurangi tingkat kompleksitas dari sebuah algoritma klasifikasi, meningkatkan akurasi dari algoritma klasifikasi tersebut, dan mampu mengetahui fitur-fitur yang paling berpengaruh terhadap tingkat akurasi.

Metode *forward selection* adalah pemodelan dimulai dari nol peubah (*empty model*), kemudian satu persatu peubah dimasukan sampai kriteria tertentu dipenuhi.

2.5 Customer Relationship Management (CRM)

CRM adalah metode dalam menjaga daya tarik pelanggan dan bagaimana tentang pilihan pelanggan, dimana merupakan aktivitas dalam merekrut pelanggan untuk menggunakan produk perusahaan dalam jangka waktu yang panjang [10].

2.6 Tinjauan Studi Penelitian Terkait

Berikut merupakan ringkasan dari penelitian terkait yang dijadikan peneliti sebagai panduan untuk penelitian yang dijelaskan pada tabel 1.

Tabel 1. Tinjauan studi penelitian terkait

No	Tahun	Judul	Penulis
1	2016	Implementasi Algoritma C4.5 Terhadap Kepuasan Pelanggan	Dhika dkk
2	2017	Penentuan Klasifikasi Pembiayaan dengan menggunakan Algoritma KNN	Ula dkk
3	2017	Analisis Kinerja Decision Tree C4.5 dalam Prediksi Potensi Pelunasan Kredit Calon Debitur	Hermanto dkk
4	2017	Implementasi Algoritma Naive Bayes Dalam Penentuan Pemberian Kredit	Rifqo dkk

3. HASIL DAN PEMBAHASAN

3.1 Pengumpulan Data

Dalam penelitian ini media pengumpulan datanya berupa data primer dan data sekunder. Data primer diperoleh melalui wawancara, observasi, dan dokumen dari instansi terkait.

- Observasi dilaksanakan pada Perusahaan Pembiayaan mulai tanggal 01 Maret 2018. Kegiatan dari observasi yang dilakukan adalah mengumpulkan data Nasabah, Agreement/Akad dan Angsuran.
- Wawancara dilakukan kepada IT head perusahaan pembiayaan, yang meliputi seputar data-data apa saja yang berkaitan dalam Customer Relationship Management (CRM).
- Dokumen, berupa beberapa contoh data *Agreement* dan nasabah dari bagian marketing

3.2 Populasi dan Sampel

Populasi adalah seluruh komponen elemen yang dapat digunakan untuk membuat beberapa kesimpulan [11]. Populasi adalah keseluruhan objek penelitian [12].

Sampel adalah bagian dari suatu populasi yang menjadi objek penelitian yang sesungguhnya [13]. Jumlah sampel yang representatif adalah 5 sampai 10 kali jumlah parameter yang digunakan [14].

Tabel 2. Populasi data.

No	Keterangan	Jumlah Data
1	Data Nasabah	1,020
2	Data Agreement / Akad	1,935
3	Data Angsuran	94,035
Jumlah		96,990

3.3 Business Understanding

Tahap pertama dari CRISP-DM adalah business understanding yaitu membuat klasifikasi dari karakteristik data nasabah/client.

- Mendapatkan data dasar dari Instansi (Data Nasabah, Pinjaman dan Angsuran)
- Melakukan Pengkodean untuk Data Nasabah dan Kode Pinjaman
- Melakukan pembobotan (kategori) untuk atribut yang telah ditentukan
- Mengklasifikasi label sesuai dengan kebutuhan perusahaan (Kolektibilitas menjadi Grade)
- Melakukan pengolahan data (Data Mining) Klasifikasi berupa Algoritma C4.5 dan K-NN menggunakan RapidMiner dengan Data Training dari Excel
- Melakukan dokumentasi terhadap hasil pengolahan data
- Mendvelop aplikasi dengan menggunakan program visual studio berdasarkan hasil kualifikasi terbaik
- Melakukan pengujian terhadap aplikasi yang telah dibuat
- Melakukan dokumentasi terhadap hasil pengujian
- Memberikan rekomendasi kepada pihak terkait berdasarkan hasil pengujian

Hasil rekomendasi ini nantinya bisa dijadikan bahan pertimbangan bagi perusahaan agar lebih fokus terhadap pelanggan.

- Mempertahankan dan memberikan perlakuan khusus bagi nasabah dengan kategori “A” berupa kemudahan proses dan pinjaman khusus.
- Mengoptimalkan nasabah dengan kategori “B” agar bisa menjadi nasabah yang berkategori “A” dengan melakukan tindakan-tindakan khusus terhadap nasabah tersebut.
- Berhati-hati terhadap nasabah dengan kategori “C” agar jangan sampai jadi kredit macet.

3.4 Data Understanding

Adapun daftar data transaksi sebagai awal pemahaman data dalam penelitian ini adalah seperti pada tabel 3.

Tabel 3. Data transaksi

Le ss ee	Less ee Type	Lsa gre e	Pe ri od	Fa cil it y Ty pe	Industr y Type	Use	Plaf ond	Banyak Tungga k	Kol ekti bili tas
HC1 012 000 1	Corpor ate	HLLS 1503 0017 0	36	LS	Perindust rian	Invest asi	840	12	5
PC1 404 006 3	Corpor ate	PLLS 1406 0026 OR1	33	LS	Jasa-jasa Dunia Usaha	Invest asi	2,349	1	5
JC11 030 001	Corpor ate	PLLS 1501 0001 OR1	8	LS	Jasa-jasa Dunia Usaha	Invest asi	261	1	5
PP1 110 000 1	Person al	PLLS 1407 0030 OR1	19	LS	Konstruks i	Invest asi	365	18	1
JP18 010 014	Person al	JMBR 1801 0014 6	1	CF	Lain-lain	Kons umsi	0,5	0	1
PP1 110 000 1	Person al	PLLS 1407 0031 OR1	19	LS	Konstruks i	Invest asi	276	25	5

3.5 Preparation data

Langkah ketiga dari teknik data Mining CRISP-DM adalah persiapan data dengan menggunakan tools RapidMiner 8.2 adapun hasilnya pada tabel 4.

Tabel 4. Data training setelah pengkodean

CLIENT	TYPE	AGR	FAC	IND	USE	AG N	PR D	AMT	QTY	COLL
151	2	1	1	3	2	1	4	4	3	5
410	2	1	1	8	2	1	4	5	1	5
106	2	4	1	8	2	1	2	3	1	5
227	1	1	1	5	2	1	3	3	3	1
598	1	1	2	11	3	1	1	1	1	1
227	1	2	1	5	2	1	3	3	3	5
227	1	3	1	5	2	1	4	4	3	5
749	2	1	1	2	2	1	4	5	1	5
721	1	2	1	2	2	1	3	3	2	5
930	2	5	1	1	2	1	5	5	1	5

Keterangan:

CLIENT: Kode Nasabah/Client

TYPE : Jenis Client (Personal/*Corporate*)

AGR : Pinjaman/Kontrak/*Agreement*

FAC : Jenis Fasilitas (LS, CF, IMBT, Murabahah)

IND : Jenis Industri

USE : Penggunaan/Peruntukan Dana

AGN : Agunan/Jaminan

PRD : Periode

AMT : Nominal Pinjaman

QTY : Banyaknya Jumlah Tunggalan

COLL : Kolektibilitas (Macet/Lancar)

Tabel 5. Kategori jenis pelanggan

Score	Kategori	Description
1	1	Personal
2	2	Corporate

Tabel 6. Kategori jenis fasilitas

Score	Kategori	Description
1	1	LS
2	2	CF
3	3	IMBT
4	4	Murabahah

Tabel 7. Kategori jenis industri

<i>Score</i>	Kategori	<i>Description</i>
1	1	Pertanian, perburuan & sarana pertanian
2	2	Pertambangan
3	3	Perindustrian
4	4	Listrik, Gas dan Air
5	5	Konstruksi
6	6	Perdagangan, restoran & hotel
7	7	Pengangkutan, pergudangan & komunikasi
8	8	Jasa-jasa Dunia Usaha
9	9	Jasa-jasa Sosial/Masyarakat
10	10	Consumer finance
99	11	Lain-lain

Tabel 8. Kategori penggunaan/peruntukan dana.

<i>Score</i>	Kategori	<i>Description</i>
1	1	Modal Kerja
2	2	Investasi
3	3	Konsumsi

Tabel 9. Kategori ketersediaan agunan

<i>Score</i>	Kategori	<i>Description</i>
1	1	Ada
2	2	Tidak Ada

Tabel 10. Kategori periode

<i>Score</i>	Kategori	<i>Description</i>
1	1	1 - 6
2	2	7 - 12
3	3	13 - 24
4	4	25 - 36
5	5	> 36

Tabel 11. Kategori nominal pinjaman

<i>Score</i>	Kategori	<i>Description</i>
1	1	<50jt
2	2	50jt-100jt
3	3	100jt-500jt
4	4	500jt-1M
5	5	1M-5M
6	6	>5M

Tabel 12. Kategori banyaknya jumlah tunggakan

Score	Kategori	Description
1	1	1-3
2	2	4-10
3	3	>10

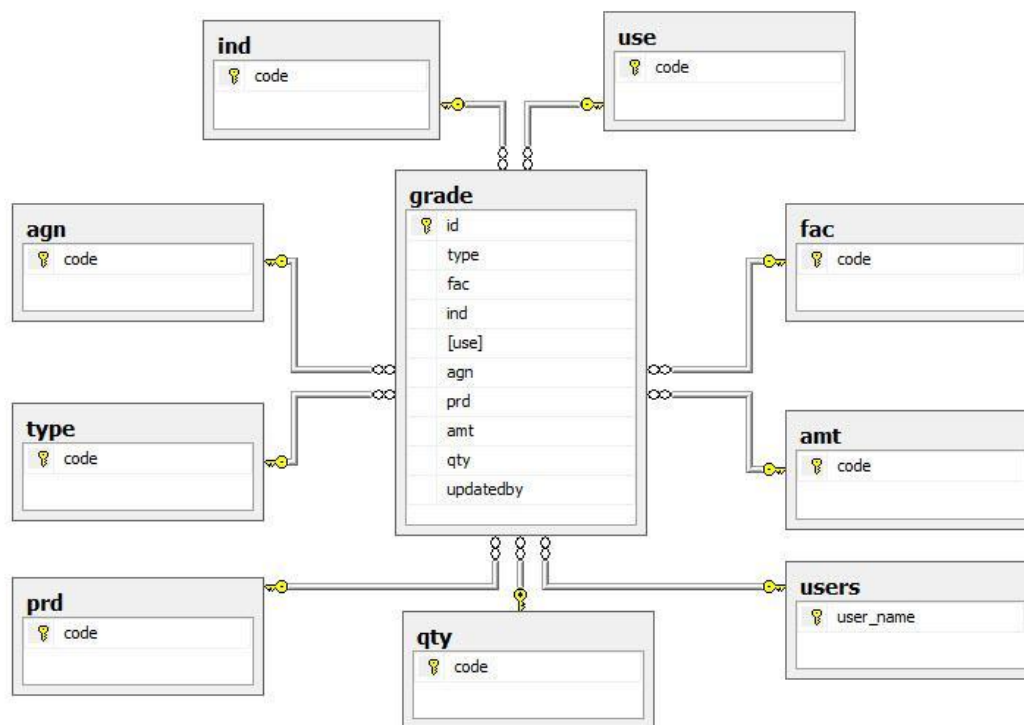
Tabel 13. Kategori kolektibilitas (*grade*)

Score	Kategori	Description
1	A	0-30 Hari
2	B	31-90 Hari
3	B	91-120 Hari
4	C	121-180 Hari
5	C	>180 Hari

Tahapan selanjutnya yaitu dilakukan proses untuk *modeling, evaluation dan deployment*.

3.6 Entity Relationship Diagram

Dalam perancangan aplikasi komparasi pelanggan ini dimana perancangan entity relationship, pada diagram gambar 2.

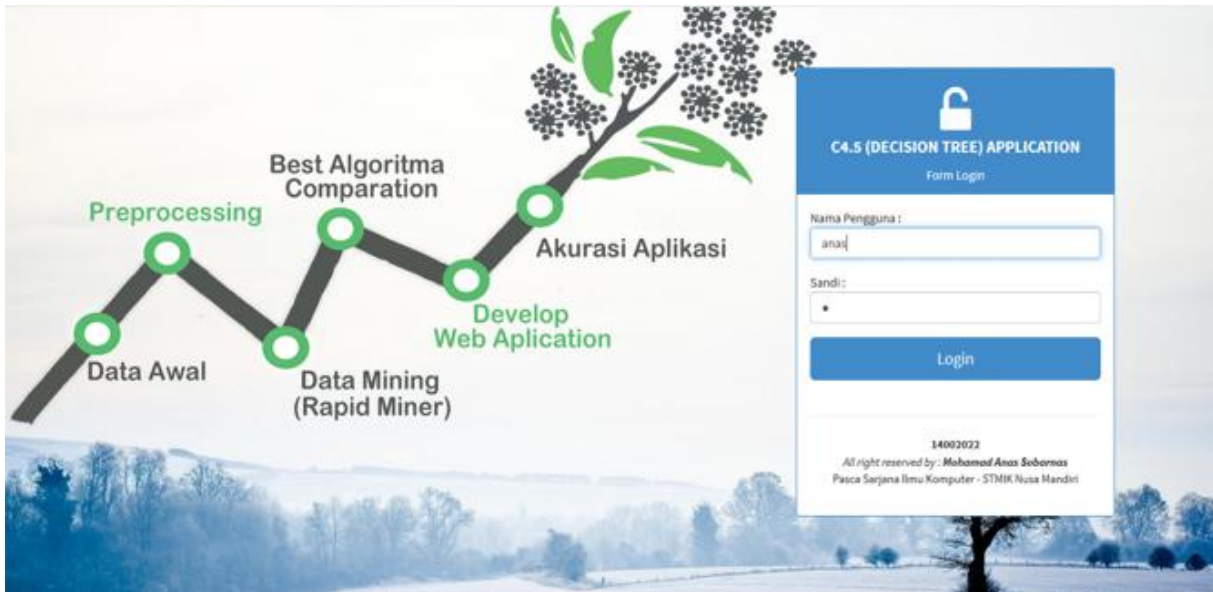


Gambar 2. Entity relationship diagram

3.7 Pengujian Aplikasi

Setelah dilakukan perancangan dan pemberian kode program maka diadakan pengujian pada aplikasi komparasi akurasi metode *k-nearest neighbor* dan C 4.5 untuk

customer relationship management pada perusahaan pembiayaan, adapun tampilan seperti pada gambar 3, 4 dan gambar 5.



Gambar 3. Form login

C4.5 (Decision Tree) Application Anas Sobarnas | Administrator

- Dashboard
- Data Set
- Aplikasi
 - List Pengujian
 - Tambah Data
 - Import Data (Excel)
- Report

Form Pengujian

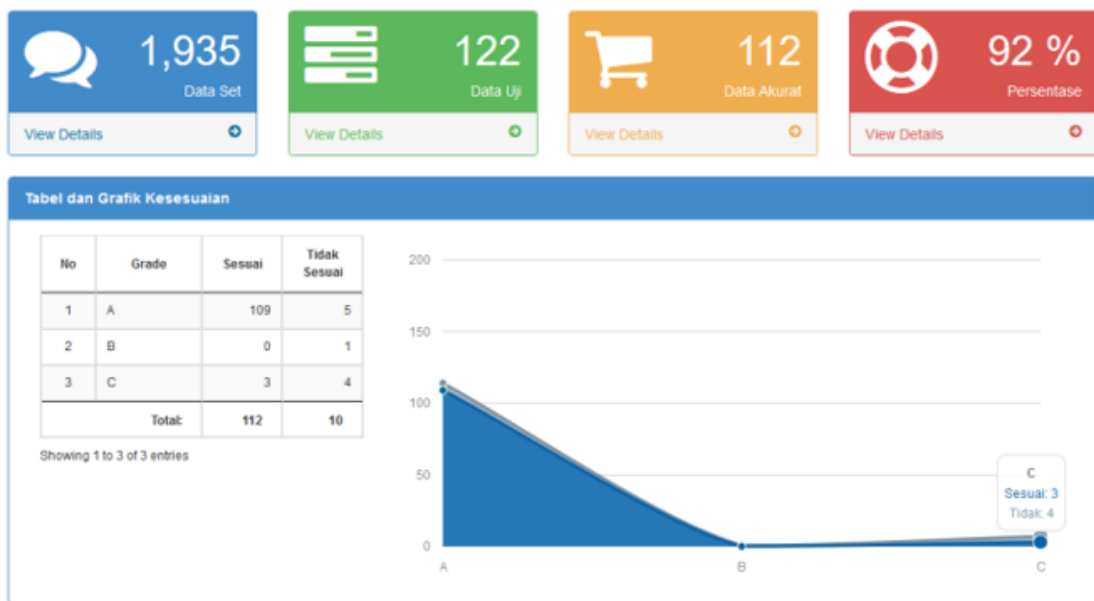
Tambah Import Clear

122 *Jml Data* 112 *Akurat* 92% *Persentase*

Show entries Search:

No	Grade (Collect) Prediksi	Grade (Collect) Fakta	Sesuai	
1	A	A	Ya	
2	A	A	Ya	
3	B	B	Ya	
4	A	A	Ya	
5	A	A	Ya	
6	A	A	Ya	
7	A	A	Ya	
8	A	A	Ya	
9	A	A	Ya	

Gambar 4. Form aplikasi proses pengujian



Gambar 5. Form aplikasi hasil pengujian

3.8 Hasil Perbandingan Algoritma

Setelah dilakukan pengujian algoritma maka dihasilkan nilai dari pengujian dua algoritma diatas seperti pada tabel 14.

Tabel 14. Perbandingan performance algoritma

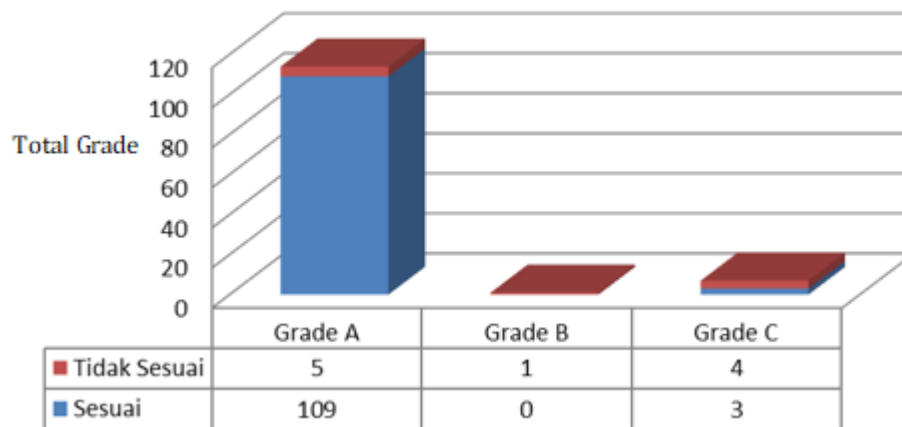
	KNN	C4.5
<i>Accuracy</i>	90.52%	93.10%
<i>Kappa</i>	0.435	0.517

Dari hasil pengujian performance kedua algoritma diatas maka menghasilkan algoritma C. 45 dengan nilai performance 93.10 % dan algoritma K-NN dengan nilai performance 90.52% selisih 2.8% sehingga algoritma C 4.5 adalah performance terbaik.

Dari hasil pengujian aplikasi, dengan jumlah data testing 122 baris data microsoft excel, didapatkan 112 data akurat dan 10 data tidak akurat maka akurasi yang didapatkan adalah 92%, yang dijelaskan pada tabel 15 dan gambar 6.

Tabel 15. Akurasi pengujian aplikasi berdasarkan grade

No	Grade	Sesuai	Tidak Sesuai
1	A	109	5
2	B	0	1
3	C	3	4
Total		112	10



Gambar 6. Grafik akurasi pengujian aplikasi berdasarkan grade

Dari hasil pengujian aplikasi berdasarkan *Grade*, dapat dilihat bahwa dari total 114 kategori A 109 Sesuai dan 5 tidak sesuai, sedangkan kategori B tidak ada data yang sesuai (0) dan 1 data tidak sesuai, dan untuk kategori C tiga sesuai dan empat tidak sesuai.

4. SIMPULAN

Berdasarkan hasil dari penelitian yang dilakukan maka bisa ditarik simpulan bahwa pembuatan aplikasi komparasi untuk klasifikasi pelanggan menggunakan teknik algoritma C 4.5 menghasilkan model klasifikasi terbaik dengan nilai 93.10 % dengan pengujian aplikasi dengan data sampel sebanyak 122 dengan akurasi klasifikasi pelanggan 92%. Maka dengan adanya aplikasi klasifikasi pelanggan berbasis web untuk produk pembiayaan ini dapat membantu manajemen dalam mengelola pelanggan dengan baik.

5. DAFTAR PUSTAKA

- [1] H. Jogiyanto, *Analisa dan Desain Sistem Informasi: Pendekatan Terstruktur Teori dan Praktik Aplikasi Bisnis*. Yogyakarta: Andi offset, 2005.
- [2] A. Susanto, *Perancangan Sistem Informasi dan Aplikasinya*. Yogyakarta: Gaya Media, 2008.
- [3] W. H. Inmon, *Building The Data Warehouse (4th ed.)*. USA: Wiley Publishing, Inc, 2005.
- [4] Kusnawi, "Pengantar solusi data mining," in *Seminar Nasional Teknologi 2007 (SNT 2007)*, 2007, pp. 1–9.
- [5] Florin Gerunescu, *Data Mining: Concept, Model, And Techniques*. Verlag Scoring Berlin Heidelberg: Springer., 2011.
- [6] L. T. E. Kusrini, *Algoritma Data Mining*. Yogyakarta: Andi offset, 2009.
- [7] F. Gerunescu, *Data Mining: Concept, Model, And Techniques*. Berlin: Verlag Scoring Berlin Heidelberg: Springer, 2011.
- [8] K. P. Han, *Data Mining concept and techinqe*. Morgan Kaufman Pubisher, 2012.
- [9] X. K. V. Wu, *The Top Ten Algorithm in Data Mining*. New York - USA: CRC Press Taylor & Francis Group, 2009.

- [10] T. S. S. & H. R. Hardiani, "Kajian Data Mining Customer Relationship Management pada Lembaga Keuangan Mikro," in *Sistem Informasi Bisnis*, 2012, pp. 181–187.
- [11] C. d. Emory, *Metode Penelitian Bisnis*. Jakarta: Erlangga, 1996.
- [12] S. Arikunto, *Prosedur Penelitian Suatu Pendekatan Praktek*. Jakarta: Rineka Cipta, 1993.
- [13] Koentjaraningrat, *Metode-metode Penelitian Masyarakat*. Jakarta: PT. Gramedia, 1993.
- [14] H. J. F. Et.al, *Multivariate Data Analysis With Reading (Fourth Edi)*. New Jersey: Prentice Hall, 1995.