PENERAPAN DATA MINING UNTUK KLASIFIKASI KELAYAKAN NASABAH DALAM PENGAJUAN KREDIT MENGGUNAKAN METODE HYBRID

Selvia Lorena BR Ginting [1], Aldi Azhar Permana [2]
[1, 2] Universitas Komputer Indonesia, Bandung
[1] selvialorena@yahoo.com, [2] azhar.aldi@yahoo.com

Abstrak - Bank merupakan sebuah lembaga yang salah satu fungsinya adalah untuk memberikan pinjaman pada nasabah sesuai jangka waktu yang telah dijanjikan dengan bunga atau tanpa bunga. Namun, setelah pihak Bank telah memberikan kreditnya atau pinjamannya muncul beberapa masalah salah satunya adalah kredit macet. Untuk menghindari hal tersebut maka diperlukan sebuah teknik yang mampu melakukan klasifikasi terhadap nasabah yang akan melakukan pengajuan kredit sehingga dapat mengurangi tingkat kredit macet. Data mining merupakan teknik yang mampu digunakan untuk menggali informasi yang tersembunyi atau pengetahuan dengan cara mencari pola-pola tertentu dari sejumlah data yang besar. Algoritma data mining yang digunakan adalah algoritma K-Nearest Neighbor yang digabungkan dengan algoritma Support Vector Machine. Hasil akhir dari penelitian ini adalah sebuah aplikasi yang diharapkan mampu menghasilkan informasi yang bermanfaat bagi Bank itu sendiri.

Kata Kunci: *Bank, Kredit, Support Vector Machine, K-Nearest Neighbor, Data Mining*

Abstract - *Bank is an institution that one of its functions is to provide loans to customers according to the promised period with interest or no interest. However, after the Bank has granted its credit or loan it appears some of the problems one of which is bad credit. To avoid this it is necessary a technique that is able to classify the customers who will make a credit application so as to reduce the level of bad debts. Data mining is a technique that can be used to extract hidden information or knowledge by searching for certain patterns of large amounts of data. The data mining algorithm used is the K-Nearest Neighbor algorithm combined with the Support Vector Machine algorithm. The end result of this study is an application that is expected to generate useful information for the Bank itself.*

Keyword : *Bank, Credit, Support Vector Machine, K-Nearest Neighbor, Data Mining*

1. PENDAHULUAN

Dengan kemajuan teknologi informasi di era modern ini, membuat kebutuhan akan informasi yang akurat sangat diperlukan sebagai acuan dalam mengambil keputsan. Dalam dunia perbankan, informasi mengenai nasabah yang melakukan kredit sangat diperlukan untuk sebagai tolak ukur dalam pemberian kredit. Namun, setelah pihak bank memberikan kredit kepada nasabah sering kali muncul beberapa masalah salah satunya ialah kredit macet.

Kredit macet atau kredit bermasalah merupakan suatu kondisi dimana pihak peminjam tidak mampu melunasi pinjaman kepada pihak yang memberi pinjaman dalam hal ini bank sesuai waktu yang telah ditentukan. Dalam lingkup kartu kredit, kredit macet termasuk kredit bermasalah adalah kondisi ketika pengguna kartu kredit telat membayar kepada pihak bank lebih dari 3 bulan [1]. Otoritas Jasa Keuangan (OJK) telah mencatat bahwa perbandingan kredit macet atau kredit bermasalah pada perbankan selama September 2016 sebesar 3,1 persen, jauh lebih tinggi dibandingkan pada saat yang sama di 2015 yang memiliki persentase sebesar 2,7 persen. Namun jika dibandingkan dengan NPL Agustus 2016 sebesar 3,22 persen, rasio kredit terancam macet mengalami penurunan tipis [2]. Untuk menghindari hal tersebut diperlukan klasifikasi terhadap nasabah yang akan melakukan pengajuan kredit. Namun dengan banyaknya jumlah nasabah akan mempersulit pihak bank dalam melakukan klasifikasi. Maka dari itu perlu sebuah teknik yang mampu melakukan klasifikasi terhadap nasabah yang akan melakukan pengajuan kredit dalam jumlah yang

banyak. Sehingga diharapkan dapat mengurangi tingkat kredit bermasalah pada bank.

Data mining merupakan sebuah teknik yang digunakan untuk mencari pola yang tersembunyi dan *knowledge* dari data yang berjumlah banyak, *data source* dapat disimpan dalam sebuah basis data website ataupun tempat penyimpanan lainnya [1]. *Data mining* juga dapat diartikan sebagai pencarian *knowledge* dalam sebuah basis data yang umumnya biasa disebut dengan istilah *Knowledge Discovery in Database* (KDD) [2]. Dengan demikian diharapkan *Data mining* dapat membantu dalam mengklasifikasikan kredit bermasalah pada nasabah.

Berdasarkan penjelasan diatas penulis mengambil judul skripsi "Penerapan Data Mining Untuk Klasifikasi Kelayakan Nasabah Dalam Pengajuan Kredit Menggunakan Metode Hybrid". Adapun metode *Hybrid* yang diusulkan pada penelitian ini adalah *Support Vector Machines* dan *K-Nearest Neighbor* yang akan di jelaskan pada bab selanjutnya.

2. TEORI PENUNJANG

2.1 Bank

Bank merupakan sebuah lembaga perantara finansial yang umumnya dibangun dengan hak untuk menerima simpanan uang dan meminjamkannya. Kata bank berakar dari bahasa Italia yaitu *banca* yang mempunyai arti tempat pertukaran uang. Sedangkan menurut undang-undang Bank adalah lembaga yang menampung uang dari masyarakat dalam bentuk tabungan dan mendistribusikannya ke masyarakat dalam bentuk pinjaman atau bentuk-bentuk yang lainnya dalam rangka menaikan taraf kehidupan masyarakat [5].

2.2 Kredit

Kata Kredit berakar dari bahasa italia yaitu *credere* yang memiliki arti kepercayaan, atau juga memiliki arti *credo* yang artinya saya percaya. Dari pengertian diatas kredit dapat diartikan sebagai kepercayaan dari pihak yang memberi pinjaman bahwa pihak yang meminjam akan mengembalikan pinjaman tersebut dengan bunga atau tanpa bunga sesuai dengan yang telah disepakati.

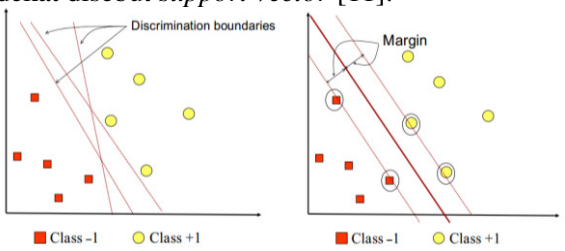
Secara umum, Kredit adalah bantuan dana dari pihak debitur dalam hal ini Bank kepada pihak kreditur perorangan atau perusahaan bahwa dana tersebut akan dikembalikan kepada pihak debitur sesuai waktu yang telah ditentukan dengan bunga atau tanpa bunga.

2.3 Data Mining

Secara umum, *data mining* merupakan sebuah teknik yang digunakan untuk mencari pola yang tersembunyi dan pengetahuan dari data yang berjumlah banyak, *data source* dapat disimpan dalam sebuah basis data website ataupun tempat penyimpanan lainnya [1]. *Data mining* juga dapat diartikan sebagai pencarian *knowledge* dalam sebuah basis data yang umumnya biasa disebut dengan istilah *Knowledge Discovery in Database* (KDD) [2].

2.4 Support Vector Machines

Support vector machines adalah sebuah metode linier dan nonlinier. Pada tahun 1992 metode s*upport vector machines* dikembangkan oleh Vladimir Vapnik bersama temannya. Ide dasarnya adalah untuk mencari *hyperplane* atau pemisah yang optimal antara satu kelas dengan kelas lainnya [1]. Metode ini dapat mengklasifikasikan data baik itu data linier atau nonlinier. Untuk mencari *hyperplane* yang optimal diantara satu kelas dengan kelas yang lainnya adalah dengan cara meghitung lebar *margin* secara masksimal. *Margin* adalah jarak dari *hyperplane* atau bidang pemisah optimal terhadap *point-point* terdekat yang berada dimasing-masing kelas. *Point-point* paling dekat disebut *support vector* [11].



Gambar 2.2 SVM berusaha menemukan hyperplane terbaik yang memisahkan kedua class

2.5 K-Nearest Neighbor

Algoritma K-Nearest Neighbor merupakan suatu metode yang menggunakan algoritma *supervised learning*[13]. Algoritma

ini pertama kali di kembangkan sekitar tahun 1950. Pengkajian K-NN didasarkan pada pembelajaran analogi, yaitu dengan membandingkan antara data *testing* dengan data *training*[1]. Klasifikasi dilakukan tidak berdasarkan model melainkan hanya berdasarkan memori. Algoritma ini bekerja berdasarkan jarak ketetanggan terdekat ke sampel baru. Untuk menghitung jarak antara atribut x1 dengan x2, maka dapat menggunakan formula *Euclidean Distance*[12]. Adapun formulanya adalah sebagai berikut:

$$d(x1, x2) = \sqrt{\sum_{i=1}^{n}(x1i - x2i)^2} \text{.....(2.1)}$$

2.6 Encoding

Encoding adalah teknik yang biasa digunakan dalam data *mining* untuk merubah data bertipe *ordinal* menjadi angka. Sebelum menemukan pola-pola pada data hal-hal yang penting dilakukan adalah persiapan data. Dalam data *mining* kebanyakan algoritma mengkhususkan input data seperti bilangan real. Ketika suatu algoritma membutuhkan sebuah inputan angka, maka teknik *encoding* harus dilakukan.

Table 2.1 contoh skema encoding.

				Disjunctive Boolean		
Original Column	*Target Column*	*Target Mean*	*Target Mean Rank*	*Red*	*Green*	*Blue*
Red	5	6	1	1	0	0
Blue	10	11	2	0	0	1
Green	20	19	3	0	1	0
Blue	12	11	2	0	0	1
Blue	11	11	2	0	0	1
Green	18	19	3	0	1	0
Red	7	6	1	1	0	0
Red	6	6	1	1	0	0
Blue	11	11	2	0	0	1

Skema encoding ditunjukan pada tabel *disjunctive boolean.* Ini adalah skema umum yang biasa digunakan ketika data tidak berupa angka [16].

2.7 Min-Max Normalization

Secara umum, kendala dalam menghitung jarak adalah jika rentang nilai pengukuran memili rentang nilai yang besar. Misalnya, atribut pendapatan dibandung dengan atribut umur, atribut pendapatan akan memiliki efek yang lebih tinggi. Untuk menghindari hal tersebut, data harus dilakukan normalisasi atau standarisasi. Tujuan dari normalisasi data adalah untuk membuat semua atribut memiliki bobot yang sama. Ada banyak teknik yang digunakan untutk melakukan normalisasi atau standarisasi data. Dalam penelitian ini digunakan teknik *min-max normalization.* Ide dasarnya adalah untuk memetakan nilai yang lama kedalam nilai yang baru sesuai inputan. Dapat dilihat dalam persamaan dibawah:

$$v`_i = \frac{v_i - a}{b-a}(mB - mA) + mA \text{.....(2.2)}$$

Dimana:

- vi adalah nilai yang akan dinormalisasi.
- a adalah nilai minimum pada sebuah kumpulan attribut.
- b adalah nilai maximum pada sebuah kumpulan attribut.
- mB adalah nilai maximum baru yang diinginkan.
- mA adalah nilai minimum baru yang diinginkan [1].

2.8 Unified Modelling Language

Unified Modelling Language merupakan bahasa yang telah menjadi standar untuk melakukan perancang sistem pada *software*, UML dapat memodelkan aplikasi yang dapat berjalan pada sistem operasi dan dapat ditulis menggunakan berbagai macam bahasa pemrograman, namun bahasa pemrograman yang sering digunakan dalam pembuatan *software* yaitu bahasa yang berorientasi objek seperti C++, JAVA, C# atau VB.NET [21].

2.9 Java

Java adalah bahasa pemrograman yang dapat dijalankan di berbagai komputer termasuk telepon genggam. Bahasa ini banyak mengadopsi sintaksis yang terdapat pada C dan C++ namun dengan sintaksis model objek yang lebih sederhana serta dukungan rutin-rutin aras bawah yang minimal. Aplikasi-aplikasi berbasis java umumnya dikompilasi ke dalam p-code (bytecode) dan dapat dijalankan pada berbagai Mesin Virtual Java (JVM). Java merupakan bahasa pemrograman yang bersifat umum/non-spesifik (general purpose), dan secara khusus didisain untuk memanfaatkan dependensi implementasi seminimal mungkin. Karena fungsionalitasnya yang memungkinkan aplikasi java mampu berjalan di beberapa platform sistem operasi yang berbeda, java dikenal pula dengan slogannya, "Tulis sekali, jalankan di mana pun". Saat ini java merupakan bahasa pemrograman yang paling populer digunakan, dan secara luas dimanfaatkan dalam pengembangan berbagai jenis perangkat lunak aplikasi ataupun aplikasi.

2.10 MySql

MySQL adalah sebuah perangkat lunak sistem manajemen basis data SQL atau DBMS yang *multithread*, *multi-user*, dengan sekitar 6 juta instalasi di seluruh dunia. MySQL AB membuat MySQL tersedia sebagai perangkat lunak gratis dibawah lisensi GNU *General Public License* (GPL), tetapi mereka juga menjual dibawah lisensi komersial untuk kasus-kasus di mana penggunaannya tidak cocok dengan penggunaan GPL.

2.11 Alpha Testing

Alpha testing merupakan simulasi atau pengujian terhadap *software* yang telah dibangun yang dilakukan pengembang, calon pengguna aplikasi, pelanggan, maupun tim uji sebelum *sofware* tersebut dipasarkan. Sedangkan pengujian Beta merupakan pengujian yang dilakuakan setelah *software* sudah digunakan oleh pengguna, untuk menentukan apakah aplikasi tersebut memenuhi kebutuhan pengguna. Pengujian Beta sering digunakan utntuk mendapatkan umpan balik dari pengguna *software* [17].

2.11.1 Metode Black-box

Pengujian *Black-box* merupakan peran yang sangat penting dalam pengujian *software*, membuat aplikasi berfungsi secara baik. Pengujian ini dilakukan berdasarkan keinginan pelanggan agar semua keinginanya terpenuhi dan keinginan yang tidak terduga dapat diidentifikasi dengan mudah, pengujian *black-box* digunakan berdasarkan persefektif penggunaan akhir untuk mangatasi kegagalan dan masukan dari pelanggan yang membuat kegagalan pada aplikasi [18].

3. ANALISIS DAN PERANCANGAN SISTEM

3.1 Analisis Algoritma

Dalam aplikasi ini algoritma yang digunakan ialah SVM dan KNN. Algoritma ini memerlukan sebuah data yang bertipe *number*. Sedangkan data yang diperoleh hampir sebagian besar memiliki tipe data ordinal kategori. Maka dari itu perlu dilakukan transformasi data pada tipe data tersebut. Sebelum melakukan transformasi data, pembuangan pada data yang tidak dibutuhkan sangatlah diperlukan karena tidak memberikan efek pada hasil mining. Dari data yang diperoleh, ada beberapa data yang akan dibuang atau tidak digunakan saat proses *mining*. Adapun data tersebut ialah id, nama dan tanggal pengajuan. Setelah dilakukan *cleaning* data atau pembuangan, hal selanjutnya yang harus dilakukan ialah transformasi data.

3.1.1 Encoding

Encoding merupakan teknik yang digunakan dalam data *mining* untuk mengubah data yang memiliki tipe ordinal *featrures* kedalam *numeric* sehingga bisa dilakukan *mining*. Adapun hasil dari teknik *encoding* adalah sebagai berikut.

Tabel 3.1 Cuplikan Data yang telah dilakukan encoding.

No.	Atribut	Nilai	Hasil
1.	Usia	< 26	1
		26 – 30	2
		31 – 35	3
		36 – 40	4
		41 – 45	5
		46 – 50	6
		51 – 55	7
		> 55	8
2.	Status Material	Duda/Janda	1
		Lajang	2
		Menikah	3

3.1.2 Min-Max Normalization

Teknik *min-max normalization* bertujuan untuk membuat semua attribut yang dipilih memiliki bobot nilai yang sama. Adapun persamaan dari pada teknik ini adalah pada persamaan 2.2.

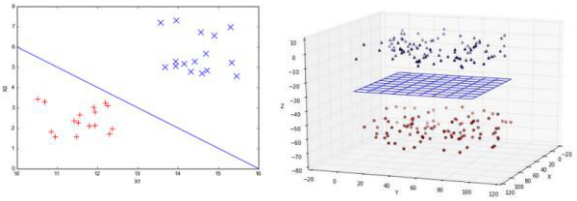
Tabel 3.2 Cuplikan Data yang telah dilakukan Min-Max Normalization.

No.	Atribut	Nilai	Hasil
1.	Usia	< 26	1
		26 – 30	1,1428
		31 – 35	1,2857
		36 – 40	1,4285
		41 – 45	1,5714
		46 – 50	1,7142
		51 – 55	1,8571
		> 55	2

2.	Status Material	Duda/Janda	1
		Lajang	1,5
		Menikah	2

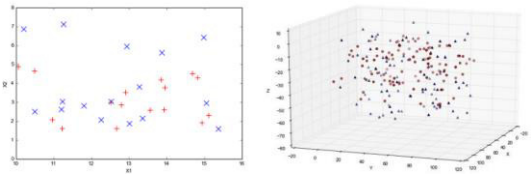
3.1.3 Support Vector Machines

Seperti yang telah di paparkan di bab sebelumnya bahwa ide dasar dari pembelajaran SVM adalah untuk menemukan pemisah antara kelas satu dengan kelas lainnya. pada gambar 2.2 sebelah kiri margin belum mendapatkan nilai yang maksimal sehingga hyperplane tidak optimal. Sedangkan pada gambar sebelah kanan margin telah mendapatkan nilai maksimal sehingga hyperplane optimal. gambar 2.2 adalah contoh kasus dengan data yang dapat dipisahkan secara linear. Kenyataannya di dalam dunia nyata data yang didapat selalu tidak linier dan sulit untuk ditemukan bidang pemisahnya.Pada gambar 3.3 adalah contoh kasus dengan data yang dapat dipisahkan secara linear. Kenyataannya di dalam dunia nyata data yang didapat selalu tidak linier dan sulit untuk ditemukan bidang pemisahnya.



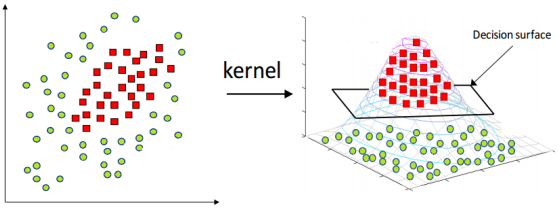
Gambar 3.2 Data Linier 2 Dimensi(kiri) dan 3 Dimensi (Kanan).

Gambar diatas adalah contoh lain dari data yang dipisahkan secara linier. Pada data linier bidang pemisah cenderung mudah untuk ditemukan pada ruang 2 dimensi ataupun 3 dimensi. Berbeda jika data tersebut adalah data yang tidak linier.



Gambar 3.3 Data Tidak Linier 2 Dimensi(kiri) dan 3 Dimensi (Kanan).

Menurut Eko Prasetyo, SVM sebenarnya hanya dapat bekerja pada data yang dipisahkan secara linear saja. Ketika data tersebut non linier maka cara yang dapat digunakan adalah menggunakan kernel pada data. Kernel dapat diartikan sebagai suatu fungsi yang memetakan ulang data pada ruang dimensi asal kedalam ruang dimensi yang lebih tinggi. Sehingga bidang pemisah dapat memisahkan data secara linear.

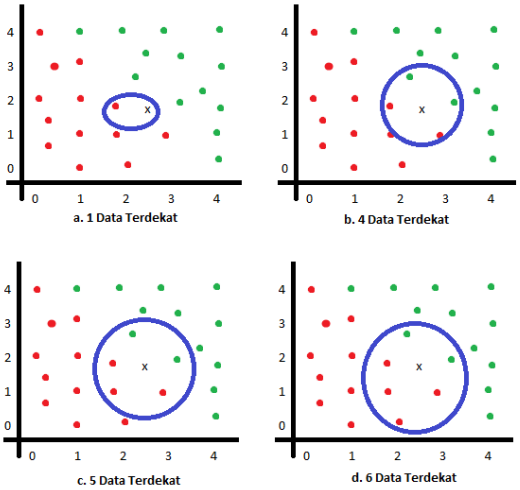


Gambar 3.4 Data Pada Ruang Dimensi Awal(kiri) dan Data Pada Ruang Dimensi Baru(Kanan).

Setelah dilakukan pemetaan pada fitur data lama kedalam fitur data baru, selanjutnya data tersebut dapat dilakukan proses pelatihan yang sama menggunakan algoritma SVM linear.

3.1.4 K-Nearest Neighbor

K-Nearst Neighbor merupakan algoritma yang bekerja berdasarkan kedekatan jarak antara data uji dengan data latih. Nilai dari parameter K pada algoritma ini berfungsi untuk menentukan jumlah data terdekat. Nilai K ini juga ditentukan oleh pengguna.



Gambar 3.5 nilai K (a) 1, (b) 4, (c) 5, (d) 6

Gambar diatas menunjukan data yang terbagi menjadi dua kelas yaitu kelas merah dan kelas hijau. Dimana kelas merah dan kelas hijau ini adalah data latih sedangkan untuk data uji adalah data yang disimbolkan dengan “x”. Terlihat pada gambar 3.6 (a) jika K bernilai satu, maka satu data latih yang terdekat dengan data uji akan dipilih sehingga kelas dari pada data uji adalah kelas merah. Pada gambar 3.6 (b) dimana K bernilai empat, kelas dari pada data uji adalah tergantung pada dekat tidaknya data uji terhadap data latih. Apabila jumlah data terdekat mayoritasnya adalah kelas merah maka

kelas dari pada data uji adalah kelas merah. Apabila jumlah data terdekat mayoritasnya adalah kelas hijau maka kelas dari pada data uji adalah kelas hijau. Sama halnya pada gambar 3.6 (c) dan (d) penentuan kelas untuk data uji adalah dengan cara mayoritas dari data latih apakah itu kelas merah atau kelas hijau.Untuk menghitung jarak antara dua titik yaitu *data testing* dan *data training* adalah dengan cara menggunakan persamaan jarak *Euclidean* yang dapat dilihat pada persamaan 2.1.

Dimana x1 adalah *data latih* dan x2 adalah data uji. Secara keseluruhan langkah-langkah dari algoritma K-NN adalah sebagai berikut:

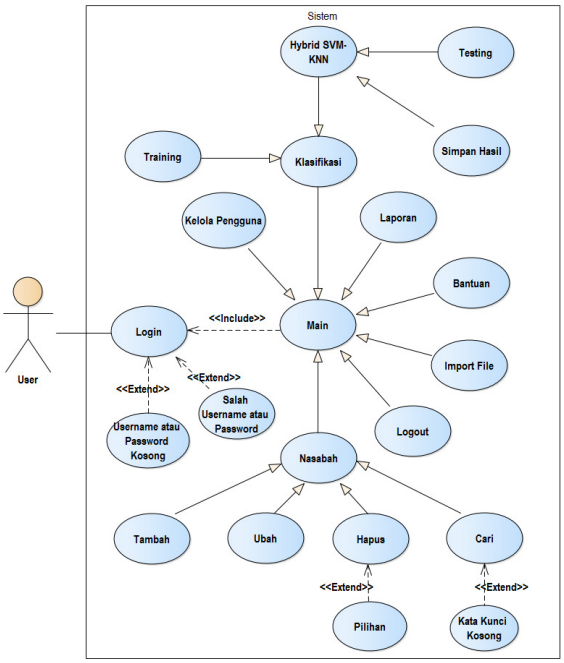
- Tentukan nilai untuk parameter K.
- Hitung jarak antara data uji dengan data latih.
- Urutkan data berdasarkan jarak data uji dengan data latih yang terkecil ke-K.
- Tentukan kelas data uji berdasarkan mayoritas data terdekat yang telah diurutkan.

3.1.4 K-Nearest Neighbor

Untuk melakukan pemodelan pada sistem bisa dilakukan dengan cara membuat pemodelan menggunakan UML (*Unified Modeling Language).*

3.2.1 Use Case Diagram

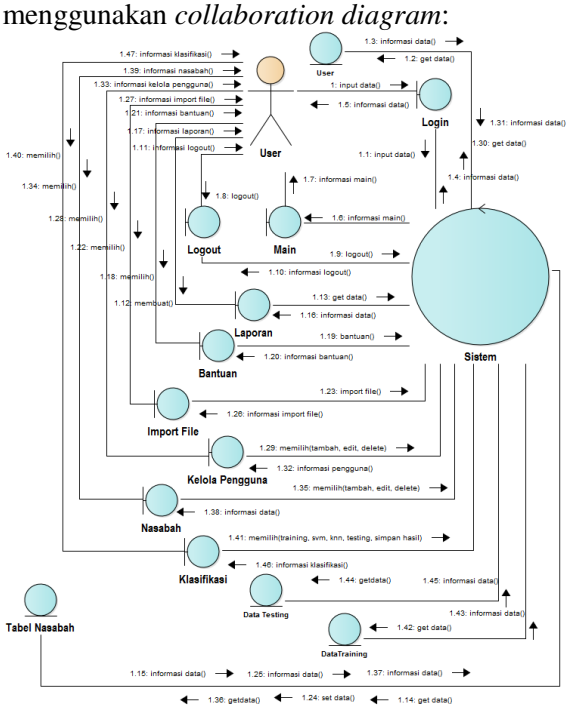
Use case diagram merupakan rangka hubungan antara aktor yang bisa berupa manusia ataupun mesin dengan sistem.



Gambar 3.6 Use Case Diagram Sistem

3.2.2 Collaboration Diagram

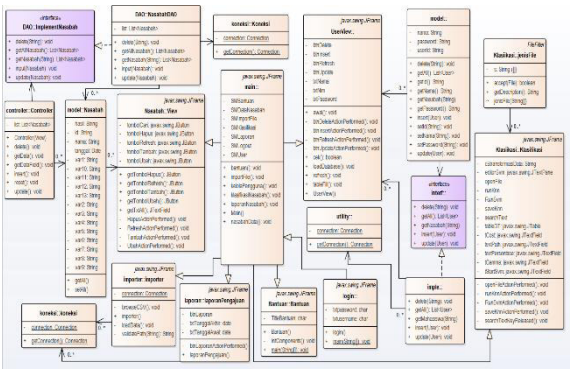
Berikut adalah gambaran sistem menggunakan *collaboration diagram*:



Gambar 3.7 Collaboration Diagram.

3.2.3 Class Diagram

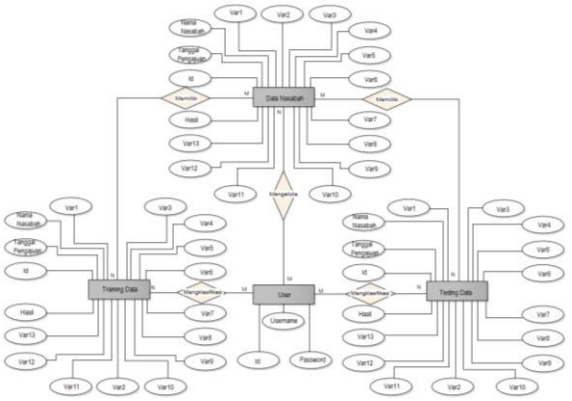
Berikut adalah gambaran sistem menggunakan *class diagram*:



Gambar 3.7 Collaboration Diagram.

3.2.3 Entity Relationship Diagram

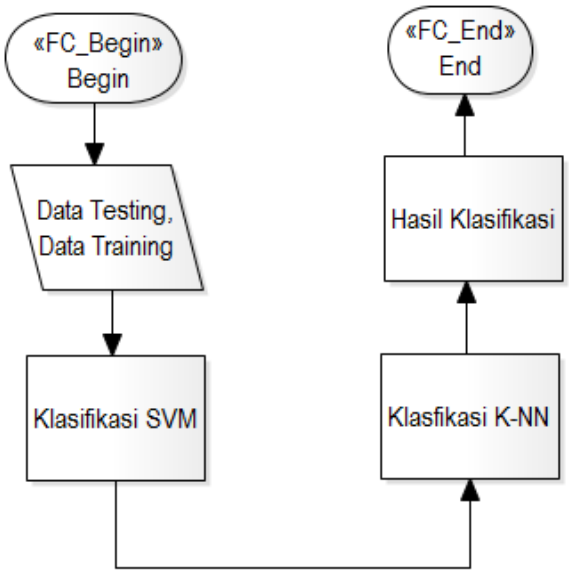
Berikut adalah gambaran sistem menggunakan *entity relationship diagram*:



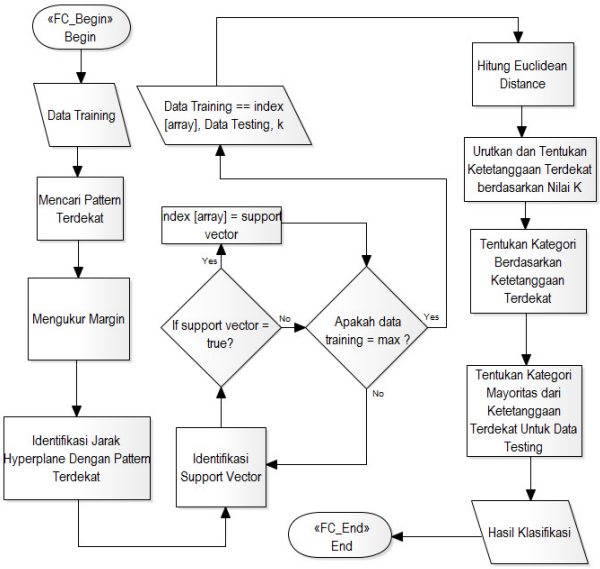
Gambar 3.8 Use Case Diagram Sistem

3.2.4 Flowchart

Dalam perancangan prosedural terdapat *flowchart* keseluruhan dan *flowchart hybrid* Adapun perancangan prosedural dapat dilihat pada gambar 3.16.



Gambar 3.9 Flowchart klasifikasi secara keseluruhan.



Gambar 3.10 Flowchart Hybrid SVM K-NN.

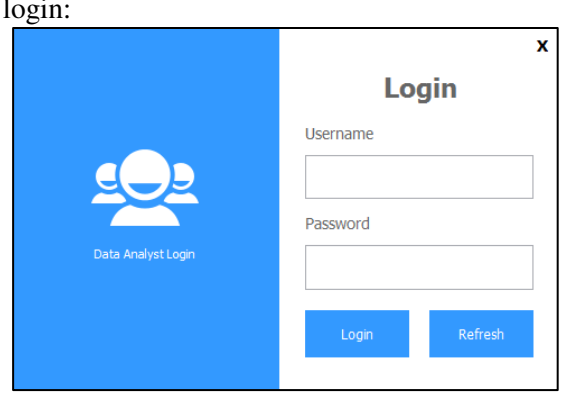
4. PENGUJIAN DAN ANALISA

4.1 Implementasi Antarmuka Aplikasi

Implementasi antarmuka merupakan penerapan hasil perancangan berupa tampilan dari aplikasi yang telah dibangun. Tampilan aplikasi ini akan digunakan oleh user. Adapun antarmuka aplikasi adalah sebagai berikut:

4.2.1 Login

Merupakan form yang pertama kali ditampilkan oleh sistem saat *user* membuka aplikasi. Berikut adalah tampilan dari menu login:

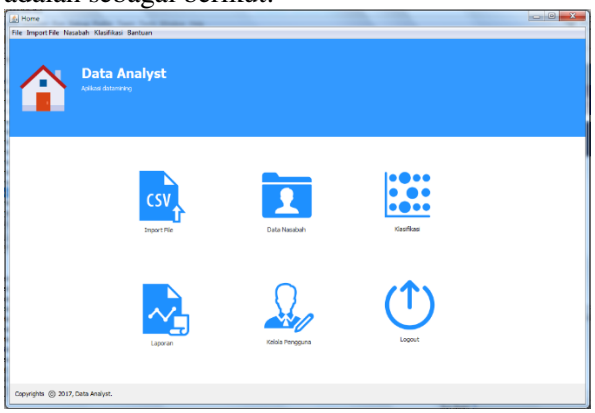


Gambar 4.1 Form Login

4.2.2 Utama

Merupakan form yang pertama kali ditampilkan oleh sistem saat *user* berhasil melakukan login. Untuk melihat form utama

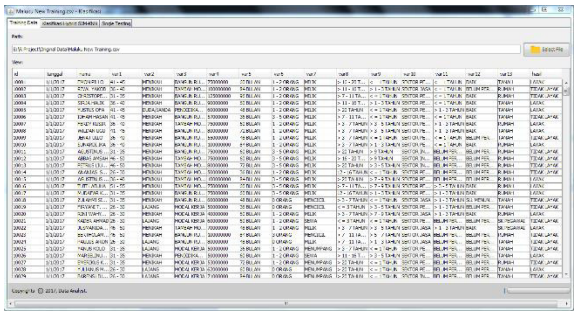
bisa dilihat pada gambar 4.2. Adapun tampilan adalah sebagai berikut:



Gambar 4.2 Form Utama

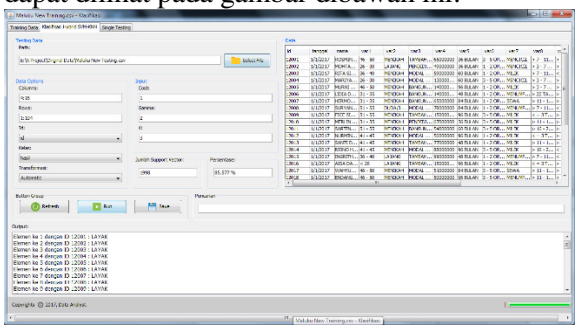
4.2.3 Klasifikasi

Merupakan menu yang digunakan oleh user untuk melakukan klasifikasi data nasabah yang melakukan pengajuan kredit. Didalam menu ini terdapat tiga tab menu dimana setiap menunya memiliki fungsi yang berbeda-beda. Adapun tab menu yang pertama adalah data training yang dapat dilihat pada gambar dibawah ini.



Gambar 4.3 Tab Menu Training Data

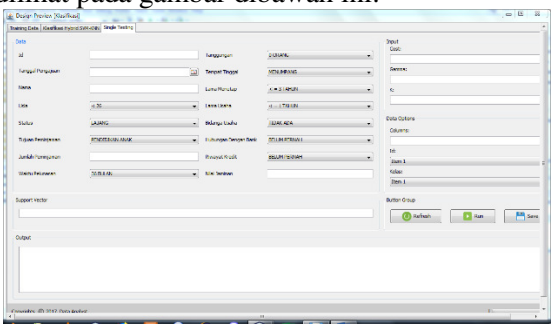
Kedua, tab menu SVM-KNN yang dapat dilihat pada gambar dibawah ini:



Gambar 4.4 Tab Menu Klasifikasi Hybrid SVM-KNN

Tab ini digunakan oleh *user* untuk menentukan layak atau tidaknya nasabah dengan jumlah data test yang banyak.

Terakhir, tab menu s*ingle test* yang dapat dilihat pada gambar dibawah ini:



Gambar 4.5 Tab Menu Single Test

4.2 Pengujian Sistem

Dalam pengujian sistem akan dilakukan pengujian pada training data dan testing data menggunakan algoritma gabungan yaitu *support vector machines* dan *k-nearest neighbor*. Adapun jumlah training data yang akan diuji adalah 2000 data sedangkan untuk jumlah testing data adalah 103 data. Adapun rencana pengunjian sistem dapat dilihat pada tabel dibawah ini:

Tabel 4.1 Rencana Pengujian Sistem Data Mining

SVM		KNN		Transformasi
Cost	Gamma	K	Data Testing	
0.1	0.02	2	103	Automatic
10	0.02	4		
0.1	0.2	8		
10	0.2	16		
0.1	2	2		
10	2	4		
0.1	20	8		
10	20	16		
2000		Default		Data Training

Sebuah sistem yang telah diuji haruslah diketahui seberapa besar tingkat keberhasilannya agar kita dapat mengambil keputusan bahwa sistem yang dibangun layak atau tidak. Dengan mengetahui jumlah data yang diklasifikasikan secara benar kita dapat

mengetahui tingkat keberhasilan dari suatu sistem. Untuk mengetahuinya kita dapat menggunakan rumus dibawah ini:

$$hasil = \frac{jumlah\ data\ yang\ benar}{jumlah\ data\ yang\ diuji} * 100\%$$

Dari delapan kali pengujian dengan diberikan nilai daripada cost, gamma dan k yang bervariasi menghasilkan *support vector* dengan persentase keberhasilan yang dapat dilihat pada tabel dibawah ini.

Tabel 4.2 Hasil Pengujian Blackbox

Parameter			Suppor Vector	Persentase (%)
Cost	Gamma	K		
0.1	0.02	2	900	57,282
10	0.02	4	650	59,223
0.1	0.2	8	1163	81,53
10	0.2	16	1214	87,379
0.1	2	20	1998	87,379
10	2	24	1997	88,35
0.1	20	32	1998	87,379
10	20	64	1997	86,408

5. KESIMPULAN

Berdasarkan hasil dari analisis, percangan dan pengujian dari aplikasi klasifikasi kelayakan nasabah menggunakan metode *Hybrid*. Kesimpulan yang didapat adalah sebagai berikut:

1. Penggabungan kedua algoritma tersebut dapat dilakukan dengan cara menggunakan total jumlah dari Support Vectors pada algoritma Support Vector Machine (SVM). Kemudian menjadikan Support Vector tersebut sebagai pengganti training data pada algortima K-Nearest Neighbor (KNN).
2. Berdasarkan hasil delapan kali pengujian diperoleh bahwa persentase keberhasilan tertinggi adalah sebesar 88,35 %. Ini mengindikasikan bahwa hasil klasifikasi menggunakan sistem dengan data uji memiliki persamaan sebesar 88,35 %. Sehingga sistem dapat dikatakan mampu untuk melakukan klasifikasi terhadap nasabah yang mengajukan kredit.
3. Faktor-faktor yang mempengaruhi hasil dari klasifikasi adalah parameter *cost* dan *gamma* pada algortima *Support Vector Machine* (SVM) dan parameter k pada *K-Nearest Neighbor* (KNN).

6. DAFTAR PUSTAKA

[1] Han, J. P. (2012). *Data Mining Concepts and Techniques 3rd Edition.* Morgan Kaufmann.

[2] Fayyad, U. P.-S. (1996). *From Data Mining to Knowledge Discovery in Databases*.

[3] https://www.cermati.com/artikel/kredit-macet-pengertian-ilustrasi-dan-efek-negatifnya (Diakses 16 Desember 2016)

[4] http://www.cnnindonesia.com/ekonomi/20161107114502-78-170819/kredit-macet-september-2016-lebih-tinggi-dibanding-tahun-lalu/ (Diakses 16 Desember 2016)

[5] http://abg01.blogspot.co.id/2014/08/pengertian-kredit-macet-penyebab-dan.html (Diakses 16 Desember 2016)

[6] https://id.wikipedia.org/wiki/Bank (Diakses 17 Desember 2016)

[7] http://www.ilmuekonomi.net/2015/10/pengertian-fungsi-sejarah-dan-jenis-jenis-bank-umum-lengkap.html (Diakses 17 Desember 2016)

[8] Larose D, T., 2005, *Discovering knowledge in data : an introduction to data mining*, Jhon Wiley & Sons Inc.

[9] https://id.wikipedia.org/wiki/Java (Diakses 16 Desember 2016)

[10] https://id.wikipedia.org/wiki/MySQL (Diakses 16 Desember 2016)

[11] A. S. Nugroho, A. B. Witarto, and D. Handoko. (2003) *Support Vector Machine: teori dan aplikasinya dalam Bioinformatika..*

[12] Rahman A.R, Nurhayati. *Implementation of Naive Bayes and K-Nearest Neighbor Algorithm for Diagnosis of Diabetes Mellitus.*

[13] Putra, Januar Adi., Akbar,Afrizal Laksita. (2016) *Klasifikasi Pengidap Diabetes Pada Perempuan Menggunakan Penggabungan Metode Support Vector Machine dan K-Nearest Neighbour.*

[14] https://kuliahinformatika.wordpress.com/2010/02/13/buku-ta-k-nearest-neighbor-knn/ (Diakses 19 Desember 2016)

[15] https://id.wikipedia.org/wiki/Microsoft_Excel

[16] Ye Nong. (2003) *The Handbook Of Datamining*.
[17] Patton Ron., 2001, *Software Testing. Indianapolis*, Sams Publishing.
[18] Ardiyansyah Feri., 2014, Implementasi Pattern Recognition pada Pengenalan Monumen-Monumen Bersejarah di Kota Bandung Menggunakan Augmented Reality Berbasis Android, j*urnal ilmiah komputer dan informatika (KOMPUTA)*, 1(8).