REPOSITORY.UB.AC.ID
UNIVERSITAS BRAWIJAYA

PENGARUH SELEKSI FITUR *INFORMATION GAIN* PADA *K-NEAREST NEIGHBOR* UNTUK KLASIFIKASI TINGKAT KELANCARAN PEMBAYARAN KREDIT KENDARAAN

SKRIPSI

Untuk memenuhi sebagian persyaratan
memperoleh gelar Sarjana Komputer

Disusun oleh:
Ulfah Mutmainnah
NIM: 155150201111362

PROGRAM STUDI TEKNIK INFORMATIKA
JURUSAN TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS BRAWIJAYA
MALANG
2019

PENGESAHAN

PENGARUH SELEKSI FITUR *INFORMATION GAIN* PADA *K-NEAREST NEIGHBOR* UNTUK KLASIFIKASI TINGKAT KELANCARAN PEMBAYARAN KREDIT KENDARAAN

SKRIPSI

Diajukan untuk memenuhi sebagian persyaratan
memperoleh gelar Sarjana Komputer

Disusun Oleh :
Ulfah Mutmainnah
NIM: 155150201111362

Skripsi ini telah diuji dan dinyatakan lulus pada
20 September 2019
Telah diperiksa dan disetujui oleh:

Pembimbing I	Pembimbing II
a.n	
Budi Darma Setiawan, S.Kom, M.Cs	Candra Dewi, S.Kom, M.Sc
NIP: 198410152014041002	NIP: 19771114 200312 2 001

Mengetahui
Ketua Jurusan Teknik Informatika

Tri Astoto Kurniawan, S.T, M.T, Ph.D
NIP: 19710518 200312 1 001

PERNYATAAN ORISINALITAS

Saya menyatakan dengan sebenar-benarnya bahwa sepanjang pengetahuan saya, di dalam naskah skripsi ini tidak terdapat karya ilmiah yang pernah diajukan oleh orang lain untuk memperoleh gelar akademik di suatu perguruan tinggi, dan tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis disitasi dalam naskah ini dan disebutkan dalam daftar referensi.

Apabila ternyata didalam naskah skripsi ini dapat dibuktikan terdapat unsur-unsur plagiasi, saya bersedia skripsi ini digugurkan dan gelar akademik yang telah saya peroleh (sarjana) dibatalkan, serta diproses sesuai dengan peraturan perundang-undangan yang berlaku (UU No. 20 Tahun 2003, Pasal 25 ayat 2 dan Pasal 70).

Malang, 22 Agustus 2019

Ulfah Mutmainnah

NIM: 155150201111362

PRAKATA

Puji syukur kehadirat Tuhan Yang Maha Esa yang telah memberikan berkah sehingga skripsi yang berjudul "Pengaruh Seleksi Fitur *Information Gain* pada *K-Nearest Neighbor* untuk Klasifikasi Tingkat Kelancaran Pembayaran Kredit Kendaraan" ini dapat terselesaikan.

Penulis menyadari bahwa dalam penyelesaian dan penyusunan skripsi ini tidak akan terwujud tanpa adanya bantuan dan dorongan dari beberapa pihak. Oleh karena itu, dengan skripsi ini penulis ingin menyampaikan rasa hormat dan terima kasih kepada:

1. Bapak Budi Darma Setiawan, S.Kom, M.Cs selaku dosen pembimbing I yang telah memberikan arahan, masukan ilmu dan bimbingan sehingga penulis dapat menyelesaikan skripsi ini.
2. Ibu Candra Dewi, S.Kom, M.Sc selaku dosen pembimbing II yang telah memberikan arahan, masukan ilmu dan bimbingan sehingga penulis dapat menyelesaikan skripsi ini.
3. Bapak Tri Astoto Kurniawan, S.T., M.T., Ph.D selaku Ketua Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Brawijaya Malang.
4. Bapak Agus Wahyu Widodo, ST. M.Cs. selaku Ketua Program Studi Informatika Fakultas Ilmu Komputer Universitas Brawijaya Malang.
5. Bapak dan Ibu dosen Fakultas Ilmu Komputer Universitas Brawijaya yang telah memberikan ilmu selama penulis melaksanakan kegiatan perkuliahan.
6. Staff dan karyawan Fakultas Ilmu Komputer yang telah membantu proses selama perkuliahan dan penulisan skripsi ini.
7. Kedua orangtua penulis yang telah memberikan nasihat, kasih sayang, perhatian dan kesabarannya di dalam membesarkan dan mendidik penulis, serta senantiasa memberikan doa dan semangat yang tiada henti-hentinya demi terselesaikannya skripsi ini.
8. Saudara penulis yang telah memberikan doa dan dukungan kepada penulis agar skripsi ini terselesaikan.
9. Teman seperjuangan dalam menyusun skripsi khususnya Rezky Amalia Hamka yang tiada henti-hentinya memberikan motivasi dan semangat kepada penulis agar skripsi ini dapat terselesaikan dan segera kembali ke kampung halaman.
10. Teman-teman "UB Telkom Squad" selaku teman-teman seperantauan yang senantiasa memberikan semangat dalam menyelesaikan skripsi ini.
11. Rekan-rekan kerja di Trans Studio Mini Malang yang telah memberikan pengalaman dan kesan selama penulis menempuh semester akhir perkuliahan dan senantiasa memberikan dukungan dalam penyelesaian skripsi ini.

Repository Universitas Brawijaya
REPOSITORY.UB.AC.ID
UNIVERSITAS BRAWIJAYA

12. Serta semua pihak yang tidak dapat penulis sebutkan satu persatu yang telah turut memberikan doa, dukungan, motivasi, dan semangat kepada penulis.

Penulis menyadari bahwa dalam penyusunan skripsi ini masih terdapat banyak kekurangan. Oleh sebab itu, penulis mengharapkan adanya saran dan kritik guna membangun untuk skripsi ini. Semoga skripsi ini dapat memberikan manfaat bagi semua pihak dan digunakan sebagaimana mestinya.

Malang, 22 Agustus 2019

Penulis

ulfahmutmainnah@student.ub.ac.id

Repository Universitas Brawijaya
REPOSITORY.UB.AC.ID
UNIVERSITAS BRAWIJAYA

REPOSITORY.UB.AC.ID
UNIVERSITAS BRAWIJAYA

ABSTRAK

Ulfah Mutmainnah, Pengaruh Seleksi Fitur *Information Gain* pada *K-Nearest Neighbor* untuk Klasifikasi Tingkat Kelancaran Pembayaran Kredit Kendaraan

Pembimbing: Budi Darma Setiawan, S.Kom, M.Cs dan Candra Dewi, S.Kom, M.Sc

Kredit macet merupakan satu diantara masalah atau risiko yang sering dihadapi oleh beberapa perusahaan penyedia jasa kredit kendaraan. Masalah tersebut berasal dari perilaku debitur yaitu tidak membayar angsuran tepat waktu. Dalam penentuan kelancaran pembayaran kredit bergantung pada analisis data-data debitur, namun dalam melakukan analisis untuk data dengan jumlah yang besar dapat memakan waktu yang lebih lama. Penelitian ini menggunakan seleksi fitur *Information Gain* dan algoritme *K-Nearest Neighbor* untuk menanggulangi masalah efektifitas dan menguji tingkat akurasi klasifikasi tingkat kelancaran pembayaran kredit kendaraan sehingga mengetahui pengaruh dari seleksi fitur. Seleksi fitur *Information Gain* yang digunakan untuk mengurangi dimensi fitur sehingga diperoleh fitur-fitur yang relevan. Fitur terpilih tersebut kemudian diproses untuk klasifikasi menggunakan algoritme *K-Nearest Neighbor*. Berdasarkan pengujian dari penelitian ini, diperoleh hasil akurasi tertinggi sebesar 94,44% pada saat pengujian dengan sebaran kelas seimbang menggunakan jumlah fitur 3 dan nilai *K*=4 sedangkan akurasi terendah diperoleh sebesar 33,33% menggunakan jumlah fitur 10 dengan nilai *K*=5 pada saat pengujian dengan sebaran kelas tidak seimbang. Fitur yang menghasilkan akurasi tertinggi yaitu pekerjaan, pendapatan dan harga On The Road (OTR). Tiga fitur tersebut merupakan fitur dengan urutan nilai *gain* terbesar dan memiliki nilai *gain* lebih dari 0,1.

Kata kunci: kredit kendaraan, seleksi fitur, *Information Gain*, klasifikasi, *K-Nearest Neighbor*

Repository Universitas Brawijaya

ABSTRACT

Ulfah Mutmainnah, ***The Effect of The Information Gain Feature Selection on K-Nearest Neighbor for The Classification of The Smoothness Rate of Auto Loan Payments***

Supervisors: Budi Darma Setiawan, S.Kom, M.Cs dan Candra Dewi, S.Kom, M.Sc

Intermittent credit is one of the problems or risks that are often faced by some auto loan service providers. The problem stems from the debtor's behavior, namely not paying the installments on time. In determining the smoothness of credit payments depends on the analysis of debtor data, but analyzing for large amounts of data can take up more time. This study uses the Information Gain feature selection and the K-Nearest Neighbor algorithm to overcome the problem of effectiveness and determine the accuracy of the classification level of the smoothness of auto loan payments so as to determine the effect of feature selection. Information Gain feature selection which is used to reduce feature dimensions so that relevant features can be obtained. The selected features are then processed for classification using the K-Nearest Neighbor algorithm. Based on testing from this study, the highest accuracy obtained is 94.44% when testing with a balanced class distribution using the number of features 3 and the value of K = 4 while the lowest accuracy is obtained at 33.33% using the number of features 10 with a value of K = 5 when testing with uneven class distribution. Features that produce the highest accuracy are jobs, income and price on the road (OTR). The three features are features with the largest order of gain values and have a gain value of more than 0.1.

Keyword: credit, feature selection, Information Gain, classification, K-Nearest Neighbor

DAFTAR ISI

PENGESAHAN ii
PERNYATAAN ORISINALITAS iii
PRAKATA iv
ABSTRAK vi
ABSTRACT vii
DAFTAR ISI viii
DAFTAR TABEL xi
DAFTAR GAMBAR xii
DAFTAR KODE PROGRAM xiii
DAFTAR LAMPIRAN xiv
BAB 1 PENDAHULUAN 1
1.1 Latar Belakang 2
1.2 Rumusan Masalah 2
1.3 Tujuan 2
1.4 Manfaat 3
1.5 Batasan Masalah 3
1.6 Sistematika Pembahasan 3
BAB 2 LANDASAN KEPUSTAKAAN 5
2.1 Kajian Pustaka 5
2.2 Kredit Kendaraan 7
2.3 Normalisasi 7
2.4 Klasifikasi 8
2.5 *K-Nearest Neighbor* (K-NN) 8
2.6 Seleksi Fitur 9
2.6.1 *Information Gain* 9
2.7 Akurasi 10
BAB 3 METODOLOGI 11
3.1 Tipe Penelitian 11
3.2 Strategi dan Rancangan Penelitian 11
3.2.1 Lokasi Penelitian 11

Repository Universitas Brawijaya
REPOSITORY.UB.AC.ID
UNIVERSITAS BRAWIJAYA

3.2.2 Metode Pengumpulan Data 11
3.2.3 Teknik Analisis Data 11
3.2.4 Perancangan Algoritme 11
3.2.5 Pengujian Sistem 12
3.2.6 Peralatan Pendukung 12
3.2.7 Penarikan Kesimpulan 13
BAB 4 PERANCANGAN 14
4.1 Perancangan Tahapan Proses 14
4.1.1 Tahapan Proses *Information Gain* 15
4.1.1.1 Tahapan Proses Hitung *Entropy* 16
4.1.1.2 Tahapan Proses Hitung *Entropy* Setiap Atribut 17
4.1.1.3 Tahapan Proses Hitung *Gain* 19
4.1.2 Tahapan Proses *K-Nearest Neighbor* 19
4.1.2.1 Tahapan Proses Hitung Jarak *Euclidean* 20
4.2 Perhitungan Manual 22
4.3 Perancangan Pengujian 32
BAB 5 IMPLEMENTASI 33
5.1 Spesifikasi Sistem 33
5.1.1 Spesifikasi Perangkat Keras 33
5.1.2 Spesifikasi Perangkat Lunak 33
5.2 Batasan Implementasi 33
5.3 Implementasi Algoritme 34
5.3.1 Implementasi Algoritme *Information Gain* 34
5.3.1.1 Implementasi Algoritme Perhitungan *Entropy* 34
5.3.1.2 Implementasi Algoritme Perhitungan *Entropy* Setiap Atribut 35
5.3.1.3 Implementasi Algoritme Pehitungan *Gain* 37
5.3.2 Implementasi Algoritme *K-Nearest Neighbor* 39

Repository Universitas Brawijaya
REPOSITORY.UB.AC.ID
UNIVERSITAS BRAWIJAYA

5.3.2.1 Implementasi Algoritme Perhitungan *Euclidean Distance* 39
5.4 Implementasi Antarmuka 40
5.4.1 Implementasi Antarmuka Data 40
5.4.1.1 Implementasi Antarmuka Data Latih 40
5.4.1.2 Implementasi Antarmuka Data Uji 40
5.4.2 Implementasi Antarmuka Proses Seleksi Fitur dan Klasifikasi 41
5.4.2.1 Implementasi Antarmuka *Information Gain* 41
5.4.2.2 Implementasi Antarmuka *K-Nearest Neighbor* 42
5.4.2.3 Implementasi Antarmuka Hasil Klasifikasi 42
5.4.3 Implementasi Antarmuka Pengujian 43
BAB 6 PENGUJIAN DAN ANALISIS 44
6.1 Pengujian 44
6.1.1 Pengujian Sebaran Kelas Seimbang pada Data Latih 44
6.1.2 Pengujian Sebaran Kelas Tidak Seimbang pada Data Latih 44
6.2 Analisis Hasil Pengujian 46
BAB 7 PENUTUP 50
7.1 Kesimpulan 50
7.2 Saran 50
DAFTAR REFERENSI 51
LAMPIRAN A DATA KREDIT KENDARAAN 53

Repository Universitas Brawijaya
REPOSITORY.UB.AC.ID
UNIVERSITAS BRAWIJAYA

DAFTAR TABEL

Tabel 2.1 Penelitian Terdahulu 6
Tabel 4.1 Data Latih 23
Tabel 4.2 Data Uji 11
Tabel 4.3 Konversi Nilai Usia 24
Tabel 4.4 Konversi Nilai Pendapatan 24
Tabel 4.5 Konversi Nilai Tanggungan 25
Tabel 4.6 Konversi Nilai Tanggungan 25
Tabel 4.7 Konversi Nilai OTR 26
Tabel 4.8 Konversi Nilai DP 26
Tabel 4.9 Konversi Nilai Tenor 27
Tabel 4.10 Konversi Nilai Angsuran 27
Tabel 4.11 Urutan Nilai Information Gain 28
Tabel 4.12 Data Latih Baru 29
Tabel 4.13 Data Uji Baru 30
Tabel 4.14 Data Latih Baru Setelah Normalisasi 30
Tabel 4.15 Data Uji Baru Setelah Normalisasi 30
Tabel 4.16 Hasil Perhitungan Jarak 31
Tabel 4.17 Hasil Pengurutan Jarak 31
Tabel 4.18 Pemilihan Data Latih Sebanyak *K* 32
Tabel 4.19 Hasil Manualisasi 32
Tabel 4.20 Perancangan Pengujian 32
Tabel 5.1 Spesifikasi Perangkat Keras 33
Tabel 5.2 Spesifikasi Perangkat Lunak 33
Tabel 6.1 Hasil Pengujian Sebaran Kelas Seimbang 44
Tabel 6.2 Hasil Pengujian Sebaran Kelas Tidak Seimbang (Skenario 1) 45
Tabel 6.3 Hasil Pengujian Sebaran Kelas Seimbang (Skenario 2) 45
Tabel 6.4 Hasil Seleksi Fitur *Information Gain* 46

Repository Universitas Brawijaya
REPOSITORY.UB.AC.ID
UNIVERSITAS BRAWIJAYA

DAFTAR GAMBAR

Gambar 3.1 Diagram Alir Proses Klasifikasi 12
Gambar 4.1 Tahapan Proses Sistem 15
Gambar 4.2 Tahapan Proses *Information Gain* 15
Gambar 4.3 Tahapan Proses Hitung *Entropy* 17
Gambar 4.4 Tahapan Proses Hitung *Entropy* Setiap Atribut 18
Gambar 4.5 Tahapan Proses Hitung *Gain* 19
Gambar 4.6 Tahapan Proses *K-Nearest Neighbor* 20
Gambar 4.7 Tahapan Proses Hitung Jarak *Euclidean* 22
Gambar 5.1 Implementasi Antarmuka Data Latih 40
Gambar 5.2 Implementasi Antarmuka Data Uji 41
Gambar 5.3 Implementasi Antarmuka *Information Gain* 42
Gambar 5.4 Implementasi Antarmuka *K-Nearest Neighbor* 42
Gambar 5.5 Implementasi Antarmuka Hasil Klasifikasi 43
Gambar 5.6 Implementasi Antarmuka Pengujian 43
Gambar 6.1 Grafik Hasil Pengujian Sebaran Kelas Seimbang 47
Gambar 6.2 Grafik Hasil Pengujian Sebaran Kelas Tidak Seimbang Skenario 1 48
Gambar 6.3 Grafik Hasil Pengujian Sebaran Kelas Tidak Seimbang Skenario 2 48

Repository Universitas Brawijaya
REPOSITORY.UB.AC.ID
UNIVERSITAS BRAWIJAYA

DAFTAR KODE PROGRAM

Kode Program 5.1 Implementasi Algoritme Perhitungan *Entropy*.....34
Kode Program 5.2 Implementasi Algoritme Perhitungan *Entropy* Setiap Atribut 36
Kode Program 5.3 Implementasi Algoritme Perhitungan *Gain*.....38
Kode Program 5.4 Implementasi Algoritme *Euclidean Distance*.....39

Repository Universitas Brawijaya
REPOSITORY.UB.AC.ID
UNIVERSITAS BRAWIJAYA

DAFTAR LAMPIRAN

Lampiran A Data Kredit Kendaraan 53

BAB 1 PENDAHULUAN

1.1 Latar Belakang

Kendaraan yang berfungsi sebagai alat transportasi sudah menjadi kebutuhan masyarakat saat ini. Dengan adanya kendaraan pribadi, dapat mempermudah kegiatan mobilitas seseorang. Kebutuhan akan kepemilikan kendaraan pribadi tidak semudah dibayangkan mengingat harga kendaraan yang tidaklah murah sehingga sebagian masyarakat tidak mampu membeli kendaraan tersebut secara tunai. Oleh karena itu, terdapat beberapa perusahaan yang dapat membantu mengatasi masalah tersebut dengan penyediaan jasa kredit kendaraan.

Kredit merupakan suatu kemampuan yang dilakukan seseorang dalam melakukan pembelian sesuatu dengan cara menerima pinjaman melalui sebuah kesepakatan pembayaran sesuai jangka waktu yang ditentukan (Astiko, 1996). Di Indonesia, terdapat banyak perusahaan yang menyediakan jasa peminjaman kredit dan tidak sedikit dari perusahaan tersebut yang mengalami berbagai masalah atau risiko. Salah satu dari risiko tersebut berasal dari perilaku debitur yang tidak membayar angsuran tepat waktu sehingga menyebabkan kredit macet. Kredit macet sangat berpengaruh kepada perusahaan karena memungkinkan terjadinya penurunan pendapatan sehingga pihak perusahaan harus lebih berhati-hati dalam menentukan calon debitur.

Dalam penentuan kelancaran pembayaran kredit bergantung pada analisis data-data debitur, namun dalam melakukan analisis untuk data dengan jumlah yang besar dapat memakan waktu yang lebih lama sehingga dibutuhkan sebuah sistem dengan metode yang dapat digunakan untuk memprediksi tingkat kelancaran pembayaran kredit kendaraan yang nantinya dapat menjadi salah satu alternatif atau sebuah acuan dari pihak perusahaan penyedia jasa kredit kendaraan tersebut dalam menentukan calon debitur.

Klasifikasi termasuk ke dalam bagian dari prediksi. Klasifikasi menggolongkan suatu data ke dalam kelompok data dimana kelompokdata tersebut kelasnya telah terdefinisi (Hsu, 2006). Terdapat beberapa penelitian mengenai klasifikasi kredit kendaraan yang salah satunya dilakukan oleh Heryono & Kardianawati (2018) dengan menggunakan algoritme Naïve Bayes yang memperoleh akurasi sebesar 65%. Terdapat beberapa metode yang dapat digunakan untuk klasifikasi, salah satunya adalah algoritme *K-Nearest Neighbor*.

Salah satu kelebihan yang dimiliki oleh algoritme K-Nearest Neighbor adalah dapat diterapkan terhadap data dengan jumlah yang besar serta memiliki banyak *noise* sehingga metode ini cukup mudah untuk diimplementasikan. Algoritme K-NN berhasil diterapkan dalam penelitian yang dilakukan oleh Leidiana (2013) untuk klasifikasi risiko kredit kendaraan bermotor yang menghasilkan nilai AUC sebesar 0,984 dan akurasi sebesar 81,46

Repository Universitas Brawijaya
REPOSITORY.UB.AC.ID
UNIVERSITAS BRAWIJAYA

%. Nilai AUC dengan rentang 0,90-1,00 merupakan klasifikasi sangat baik (Gorunescu, 2011). Namun salah satu kekurangan dari K-NN yaitu pemilihan atribut terbaik (Bahri & Maliki, 2012). Pemilihan atribut terbaik dapat dilakukan dengan cara seleksi fitur.

Seleksi fitur dapat dilakukan sebagai proses *pre-processing* klasifikasi dengan menghilangkan fitur yang tidak memiliki pengaruh besar. Fitur yang tidak relevan dan jumlahnya banyak dapat mempengaruhi masalah efisiensi sehingga seleksi fitur dapat membantu mengenali keseluruhan data tersebut (Poonguzhali, et al., 2014). Proses seleksi fitur menutupi kekurangan K-NN karena dapat mengurangi dimensi data sehingga dapat menghasilkan akurasi yang lebih baik. Penelitian yang dilakukan oleh Ivandari, et al., (2017) membuktikan bahwa penerapan *Information Gain* untuk seleksi fitur dapat mempengaruhi performa algoritme K-NN dalam klasifikasi persetujuan kredit. Performa algoritme K-NN naik hingga lebih dari 3% dengan menggunakan seleksi fitur.

Bersumber pada penelitian-penelitian yang telah dilakukan tersebut, diketahui bahwa seleksi fitur dengan metode *Information Gain* dan algoritme *K-Nearest Neighbor* dapat diterapkan untuk mengklasifikasi data kredit kendaraan sehingga dalam penelitian ini, kedua metode tersebut diharapkan dapat mengetahui hasil akurasi klasifikasi serta dapat mengidentifikasi fitur atau atribut yang berpengaruh dalam data kredit kendaraan.

1.2 Rumusan Masalah

Berdasarkan pada permasalahan yang telah dijabarkan di sub bab sebelumnya, maka permasalahan tersebut dirumuskan sebagai berikut:

1. Berapa nilai akurasi pada klasifikasi tingkat kelancaran pembayaran kredit kendaraan menggunakan seleksi fitur *information gain* dan algoritme *k-nearest neighbor*?
2. Fitur apa yang dapat menghasilkan nilai akurasi tertinggi dan memberikan pengaruh pada klasifikasi tingkat kelancaran pembayaran kredit kendaraan menggunakan algoritme *k-nearest neighbor*?

1.3 Tujuan

Sebagai jawaban dari permasalahan yang telah diuraikan, maka diharapkan tercapainya tujuan dalam penelitian ini sebagai berikut:

1. Mengetahui akurasi klasifikasi tingkat kelancaran pembayaran kredit kendaraan menggunakan algoritme *k-nearest neighbor*.
2. Mengetahui fitur yang menghasilkan nilai akurasi tertinggi serta mengetahui pengaruh pada klasifikasi tingkat kelancaran pembayaran kredit kendaraan menggunakan algoritme *k-nearest neighbor*.

1.4 Manfaat

Manfaat yang didapatkan dalam penelitian ini yaitu dapat memudahkan proses analisis kredit dengan jumlah data yang besar melalui proses klasifikasi serta akurasi yang dihasilkan lebih baik dengan adanya seleksi fitur.

1.5 Batasan Masalah

Agar tercapainya tujuan dan maksud yang sesungguhnya dari penelitian ini, maka ditentukan batasan sebagai berikut:

1. Dataset yang digunakan dalam penelitian ini didapatkan dari salah satu lembaga finansial penyedia jasa kredit kendaraan di kota Makassar.
2. Data yang digunakan sebagai fitur/atribut sebanyak 10, yaitu jenis kelamin, usia, status, pekerjaan, pendapatan, tanggungan, harga *On The Road* (OTR), *Down Payment* (DP), tenor, serta angsuran.

1.6 Sistematika Pembahasan

Sistematika pembahasan dijabarkan melalui beberapa bab yang secara keseluruhan mewakili isi dari naskah skripsi ini. Adapun pembagiannya adalah sebagai berikut:

BAB 1 PENDAHULUAN

Bab ini menjelaskan mengenai persoalan yang sedang terjadi yang diuraikan ke dalam latar belakang dilakukannya penelitian, rumusan masalah dan batasan masalah yang diangkat pada penelitian, tujuan, manfaat dari diterapkannya metode yang digunakan serta sistematika pembahasan.

BAB 2 LANDASAN KEPUSTAKAAN

Bab ini menjelaskan mengenai dasar teori atau referensi terkait metode yang digunakan dalam pemahaman permasalahan penelitian. Adapun teori-teori atau referensi tersebut tentang kredit, klasifikasi, seleksi fitur, *Information Gain*, dan *K-Nearest Neighbor*.

BAB 3 METODOLOGI

Bab ini menjelaskan uraian mengenai metode yang digunakan dalam proses penyelesaian penelitian ini yang meliputi tipe penelitian, lokasi dilaksanakannya penelitian, cara mengumpulkan data, teknik dalam analisis data, perancangan algoritme dan pengujian, serta penarikan kesimpulan.

BAB 4 PERANCANGAN

Bab ini menjelaskan berkenaan dengan proses atau uraian langkah dalam penerapan metode yang digunakan pada penelitian ini serta perancangan dalam pengujian.

Repository Universitas Brawijaya
REPOSITORY.UB.AC.ID
UNIVERSITAS BRAWIJAYA

BAB 5 IMPLEMENTASI

Bab ini menjelaskan mengenai rangkaian diterapkannya metode seleksi fitur *Information Gain* dan algoritme *K-Nearest Neighbor* berdasarkan perancangan yang telah dibuat di bab sebelumnya.

BAB 6 PENGUJIAN DAN ANALISIS

Bab ini menjelaskan mengenai cara pengujian dan menganalisa tingkat akurasi dari metode yang digunakan.

BAB 7 PENUTUP

Bab ini menjelaskan mengenai kesimpulan dari hasil pengujian penilitian yang telah dilakukan dan saran dari penulis bagi pembuatan penelitian terkait yang selanjutnya.

Repository Universitas Brawijaya
REPOSITORY.UB.AC.ID
UNIVERSITAS BRAWIJAYA

BAB 2 LANDASAN KEPUSTAKAAN

2.1 Kajian Pustaka

Kajian pustaka dilakukan dengan cara mengumpulkan penelitian-penelitian yang berkaitan dengan metode yang akan diterapkan dalam penelitian ini dan dapat dijadikan sebagai sumber acuan dalam menyelesaikan penelitian ini.

Penelitian pertama yang membahas pemanfaatan algoritme *K-Nearest Neighbor* pada data pengguna jasa pembiayaan kredit kendaraan dilakukan oleh Henny Leidiyana pada tahun 2013. Penelitian ini menggunakan 481 data dan fitur sejumlah 14. Nilai akurasi yang didapatkan melalui pengujian sebesar 81,46% dan nilai AUC sebesar 0,984 dan termasuk dalam klasifikasi sangat baik.

Penelitian selanjutnya dilakukan oleh Heryono dan Acun Kardianawati pada tahun 2018 yang bertujuan untuk memprediksi layak tidaknya seseorang diberi kredit dengan menggunakan metode *Naïve Bayes*. Data yang digunakan berjumlah 115 data dengan atribut sejumlah 6 yaitu jenis kelamin, usia, status, pekerjaan, penghasilan, dan masa pembayaran. Pengujian dilakukan menggunakan 28 data yang menghasilkan akurasi sebesar 65,00%.

Penelitian berikutnya masih membahas mengenai kredit dengan penerapan seleksi fitur *Information Gain* dan metode klasifikasi *K-Nearest Neighbor*. Penelitian ini dilakukan oleh Ivandari, et al. pada tahun 2017. Dalam penelitian ini menggunakan 766 data kredit dengan kelas kredit lancar sebanyak 210 dan sisanya termasuk dalam kelas kredit macet. Terdapat 14 fitur yang akan diberikan bobot menggunakan Information Gain untuk selanjutnya diklasifikasi. Dari pengujian klasifikasi yang dilakukan didapatkan akurasi tertinggi pada data UCI sebesar 74,93% sebelum menerapkan seleksi fitur dan akurasinya meningkat menjadi 82,46% dengan hanya menggunakan 5 fitur, sedangkan pada data lokal dihasilkan akurasi tertinggi sebesar 91,52% sebelum menerapkan seleksi fitur dan akurasinya meningkat menjadi 94,78% dengan hanya menggunakan 6 fitur.

Penelitian selanjutnya yang menerapkan seleksi fitur *Information Gain* dilakukan oleh Aini, et al. pada tahun 2018. Pada penelitian tersebut menerapkan *Information Gain* pada untuk menyeleksi fitur pada data penyakit jantung dengan menerapkan algoritme *K-Nearest Neighbor* lalu mengkombinasikan dengan algoritme *Naïve Bayes* untuk mengklasifikasikan tingkat risikonya. Akurasi paling tinggi yang diperoleh dengan menerapkan seleksi fitur sebesar 92,31% sedangkan akurasi dengan tanpa seleksi fitur sebesar 80,77%.

Penelitian berikutnya yang menerapkan metode *Information Gain* untuk prediksi menggunakan *K-Nearest Neighbor* yang dilakukan oleh Arifin pada tahun 2015. Penelitian dilakukan untuk memprediksi peluang pelanggan untuk meninggalkan jasa telekomunikasi yang mereka gunakan. Hasil penelitian tersebut menunjukkan kenaikan akurasi sebesar 1,7% dengan menggunakan seleksi fitur *Information Gain* dibandingkan dengan hanya menggunakan *K-Nearest Neighbor*.

Adapun beberapa jurnal ataupun penelitian terkait tersebut ditunjukkan pada Tabel 2.1.

Tabel 2.1 Penelitian Terdahulu

No.	Judul	Metode	Keterangan
1.	"Penerapan Algoritma *K-Nearest Neighbor* Untuk Penentuan Resiko Kredit Kepemilikan Kendaraan Bermotor" (Leidiana, 2013)	*K-Nearest Neighbor*	Akurasi yang didapatkan melalui pengujian sebesar 81,46% dan nilai AUC sebesar 0,984 dan termasuk dalam klasifikasi sangat baik.
2.	"Implementasi Metode Naive Bayes Untuk Klasifikasi Kredit Motor" (Heryono & Kardianawati, 2018)	*Naïve Bayes*	Akurasi sebesar 65,00%
3.	"*Data Attribute Selection with Information Gain to Improve Credit Approval Classification Performance using K-Nearest Neighbor Algorithm*" (Ivandari, et al., 2017)	*Information Gain* dan *K-Nearest Neighbor*	Terjadi peningkatan akurasi pada klasifikasi dengan penambahan seleksi fitur pada data UCI sebesar 7,53% menjadi 82,46% dan pada data lokal meningkat sebesar 3,26% menjadi 94,78%
4.	"Seleksi Fitur *Information Gain* untuk Klasifikasi Penyakit Jantung Menggunakan Kombinasi Metode *K-Nearest Neighbor* dan *Naïve Bayes*" (Aini, et al., 2018)	Gabungan antara metode *K-Nearest Neighbor* dengan *Naïve Bayes*	Akurasi tertinggi yang diperoleh sebesar 92,31%
5.	"IG-KNN untuk Prediksi Customer Churn Telekomunikasi" (Arifin, 2015)	*Information Gain* dan *K-Nearest Neighbor*	Hasil akurasi meningkat 1,7% dengan menggunakan algoritme IG-KNN dibandingkan hanya dengan menggunakan KNN saja

2.2 Kredit Kendaraan

Kredit merupakan suatu kegiatan yang dilakukan dengan memberikan barang, jasa, ataupun uang dari satu pihak yang disebut kreditor kepada pihak lain yang disebut debitur/nasabah melalui sebuah kesepakatan pembayaran kembali sesuai jangka waktu yang telah ditentukan (Rivai, 2006). Sehingga dapat disimpulkan bahwa kredit kendaraan berarti memberikan kendaraan kepada debitur/nasabah yang kemudian debitur tersebut akan membayar angsuran sesuai kesepakatan sejumlah harga kendaraan yang diberikan. Peningkatan permintaan kredit setiap tahunnya menjadi tantangan bagi perusahaan penyedia kredit untuk memperkecil masalah atau risiko yang akan dihadapi dengan cara melakukan analisis kredit.

Menurut Rivai (2006), analisis kredit adalah suatu aktivitas yang dilakukan untuk memperoleh informasi mengenai kelayakan dari suatu persoalan kredit. Dengan adanya analisis kredit, diperoleh salah satu permasalahan yang sering terjadi yaitu apakah nasabah nasabah dapat membayar angsuran tepat waktu sehingga tidak menyebabkan kredit macet. Dalam melakukan analisis kredit, diperlukan beberapa data nasabah yang nantinya akan digunakan sebagai bahan analisis.

2.3 Normalisasi

Salah satu tahap dalam penelitian adalah *prepocessing data*. Pada penelitian ini, dilakukan tahap transformasi data sehingga data tersebut dapat membentuk pola tertentu yang dapat dijadikan sebagai informasi (Kamber & Han, 2006). Salah satu metode tranformasi data yang dilakukan dalam penelitian ini adalah normalisasi. Normalisasi merupakan suatu proses yang dilakukan pada nilai atribut dari data agar nilai tersebut terletak pada rentang tertentu yang lebih kecil.

Salah satu teknik normalisasi data yang paling sering digunakan yaitu normalisasi *min-max*. Normalisasi *min-max* merupakan metode yang dilakukan dengan cara mengubah skala data menjadi range baru. Dengan dilakukannya normalisasi *min-max*, nilai-nilai yang akan digunakan menjadi seimbang karena berada pada rentang yang sama. Atribut yang dinormalisasi nilainya akan berada pada rentang 0 sampai 1 (Junaedi, et al., 2011).

Berikut merupakan persamaan dalam normalisasi *min-max*:

$$X' = \frac{X - \min(X)}{\max(X) - \min(X)} \quad (2.1)$$

Dimana:

X' : Nilai data baru dari hasil normalisasi *min-max*
$\max(X)$: Nilai maksimal dari data
$\min(X)$: Nilai minimal dari data

2.4 Klasifikasi

Klasifikasi termasuk ke dalam proses prediksi yang termasuk ke dalam penerapan ilmu data mining. Klasifikasi merupakan suatu runtutan perubahan untuk menemukan suatu pola yang mewakili dari sekelompok data yang memiliki label kelas yang telah terdefinisi sehingga nantinya data baru dapat ditentukan kelasnya melalui pola yang telah terbentuk tersebut (Kamber & Han, 2006).

Dua tahap utama dalam klasifikasi terdiri atas pembelajaran yang dilakukan untuk menganalisis data latih lalu direpresentasikan menjadi sebuah aturan klasifikasi sebagai tahap pertama. Proses kedua yaitu klasifikasi yang dilakukan untuk memprediksi hasil dari aturan klasifikasi yang telah diperoleh dari proses pertama.

2.5 *K-Nearest Neighbor* (K-NN)

K-Nearest Neighbor (K-NN) termasuk dalam metode *lazy learning* atau berbasis pembelajaran. *K-Nearest Neighbor* (K-NN) digunakan untuk klasifikasi data baru dengan cara mencari *k* kelompok dari jarak yang paling dekat antara data baru tersebut dengan beberapa data tetangga (Santoso, 2007).

Langkah-langkah dalam algoritme *K-Nearest Neighbor* (K-NN) dapat dijabarkan sebagai berikut (Santoso, 2007):

1. Menentukan nilai dari *K*. *K* merupakan jumlah tetangga yang jaraknya paling dekat dengan latih.
2. Menghitung jarak kedekatan semua data latih ke data uji.
3. Mengurutkan data yang memiliki jarak dari yang terkecil ke yang terbesar.
4. Mengambil data latih terdekat sejumlah *K* yang sudah ditentukan sebelumnya.
5. Periksa label kelas dari data yang telah diurutkan.
6. Menentukan kelas sesuai dengan label kelas yang memiliki frekuensi paling banyak.

Pada penelitian ini menggunakan *euclidean distance* yang merupakan cara menghitung jarak yang cukup populer dengan persamaan sebagai berikut (Bramer, 2007):

$$D(x,y) = \sqrt{\sum_i^n (x_i{}^2 - y_i{}^2)} \quad (2.5)$$

Dimana D merupakan jarak kedekatan. x merepresentasikan data latih dan y merepresentasikan data uji.

Penelitian ini menggunakan 10 fitur yang 7 diantara fitur tersebut berupa data numerik dan 3 fitur sisanya berupa data nominal atau yang biasa dikenal dengan data kategori. Untuk menghitung jarak data nominal menggunakan rumus sebagai berikut (Larose, 2005):

$$D_i = \begin{cases} 0 & jika\ x_i = y_i, \\ 1 & lainnya \end{cases} \quad (2.6)$$

Repository Universitas Brawijaya
REPOSITORY.UB.AC.ID
UNIVERSITAS BRAWIJAYA

Fungsi tersebut menyatakan jika data *training* dan data *testing* sama berarti jaraknya bernilai 0, sebaliknya jika data latih dan data uji berbeda maka nilai jaraknya bernilai 1.

2.6 Seleksi Fitur

Pada tahap *preprocessing* data, terdapat salah satu proses yang dinamakan seleksi fitur. Seleksi fitur merupakan salah satu metode untuk memperkecil atau mengurangi dimensi atribut sehingga menghasilkan atribut yang dianggap relevan untuk selanjutnya diproses dalam *data mining*. Jumlah atribut yang banyak dan jika atribut tersebut tidak relevan akan mempengaruhi waktu dalam proses klasifikasi dan nilai akurasi yang dihasilkan dalam algoritme klasifikasi yang digunakan (Witten, et al., 2011). Oleh karena itu, seleksi fitur diperlukan untuk mengurangi dimensi atribut tersebut sehingga memungkinkan meningkatnya hasil akurasi dari suatu algoritme.

Adapun metode seleksi fitur yang banyak digunakan oleh peneliti diantaranya adalah *Information Gain* (IG), *Backward Elimination*, dan *Forward Selection.* Dalam penelitian ini akan menggunakan seleksi fitur *Information Gain* (IG).

2.6.1 *Information Gain*

Information Gain merupakan salah satu metode dalam proses seleksi fitur yang dilakukan dengan cara melakukan perangkingan nilai bobot atribut (Chormunge & Jena, 2016). *Information Gain* dapat membantu menghasilkan atribut-atribut yang relevan terhadap kelas target karena setiap atribut memiliki nilai dan dapat dipilih yang terbaik. Penentuan atribut relevan diperoleh melalui nilai *entropy*. *Entropy* merupakan suatu takaran ketidakpastian kelas dengan menggunakan peluang kemunculan dari atribut tertentu (Shaltout, et al., 2014). Pengukuran nilai *entropy* tersebut menjadi awal dari patokan untuk menentukan atribut-atribut apa saja yang mencukupi kriteria bedasarkan pembobotan dan akan digunakan dalam proses klasifikasi sebuah algoritme.

Berikut merupakan persamaan dalam perhitungan *Information Gain* (Azhagusundari & Thanamani, 2013):

$$Entropy(S) = \sum_{i=1}^{c} -P_i log_2 P_i \quad (2.2)$$

Dimana c merupakan jumlah dari kelas klasifikasi dan P_i adalah jumlah peluang pada kelas i.

$$Entropy(A) = \sum_{j=1}^{v} \frac{|S_j|}{|S|} \times entropy(S_j) \quad (2.3)$$

$$Gain(A) = Entropy(S) - Entropy(A) \quad (2.4)$$

Dimana:

A : atribut

v : jumlah partisi atribut

$|S_j|$: jumlah sampel pada partisi ke j

$|S|$: jumlah seluruh sampel data

$entropy(S_j)$: total *entropy* di dalam partisi

$Entropy(S)$: total *entropy*

$Entropy(A)$: entropy untuk atribut

$Gain(A)$: nilai gain untuk atribut

2.7 Akurasi

Akurasi merupakan satuan hasil yang diukur berdasarkan hasil dari uji coba kinerja sistem yang telah dibuat dengan menerapkan metode tertentu. Nilai akurasi diperoleh dengan membandingkan hasil pengujian antara nilai prediksi sistem dengan nilai sebenarnya. Perhitungan akurasi ditunjukkan pada persamaan 2.7.

$$Akurasi\ (\%) = \frac{\sum data\ uji\ benar}{\sum total\ data\ uji} x100\% \quad (2.7)$$

BAB 3 METODOLOGI

3.1 Tipe Penelitian

Tipe penelitian yang digunakan dalam penelitian ini adalah non-implementatif analitik. Penelitian nonimplementatif merupakan kegiatan yang dilakukan untuk mengamati serta mendapatkan informasi berdasarkan situasi atau fenomena yang ada. Studi kasus pada penelitian ini adalah data kredit kendaraan macet. Pendekatan yang diterapkan dalam penelitian ini yaitu analitik (*analytical/explanatory*).

Metode analitik merupakan pendekatan yang menguraikan secara terperinci hubungan antara bagian-bagian dari objek yang sedang diteliti, dalam kasus ini kredit kendaraan merupakan objek penelitian. Pendekatan ini akan menghasilkan jawaban dari permasalahan yang telah dirumuskan sebelumnya.

3.2 Strategi dan Rancangan Penelitian

3.2.1 Lokasi Penelitian

Penelitian yang membahas mengenai klasifikasi tingkat kelancaran pembayaran kredit kendaraan ini dilaksanakan di Laboratorium Komputasi Cerdas Fakultas Ilmu Komputer Universitas Brawijaya.

3.2.2 Metode Pengumpulan Data

Data yang digunakan dalam penelitian ini yaitu data sekunder yang berasal dari perusahaan X sebagai penyedia jasa kredit. Data tersebut diperoleh dari pemangku kepentingan yang berada di perusahaan penyedia jasa kredit tersebut. Data yang didapatkan meliputi semua aspek dari pengguna jasa kredit yang nantinya akan digunakan sebagai fitur, yaitu jenis kelamin, usia, status perkawinan, pekerjaan, pendapatan, tanggungan, harga On The Road (OTR), DP tenor, angsuran serta keterangan (lancar, diragukan, dan macet). Total data yang diperoleh sejumlah 240 data.

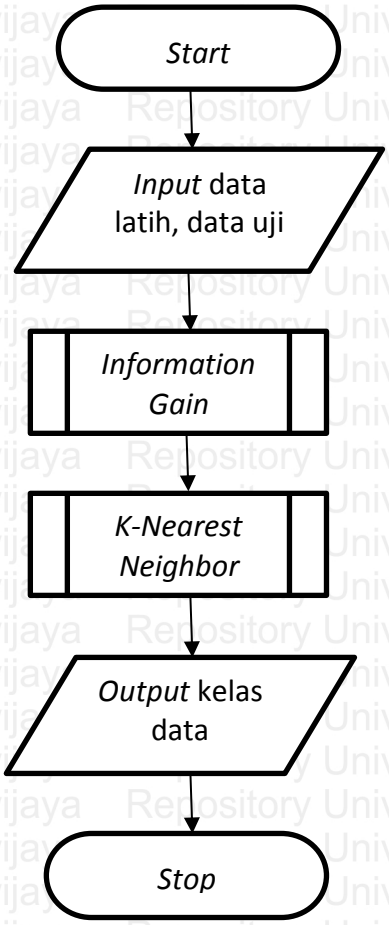
3.2.3 Teknik Analisis Data

Penilitian ini menggunakan teknik analisis data berupa analisa data kuantitatif yang dilakukan dengan cara mengelompokkan data sesuai atributnya di dalam tabel dan kemudian mengkonversi data numerik menjadi katagorikal untuk diproses melalui *information gain*. Setelah proses *information gain* data kembali dikonversi menjadi data numerik agar dapat diklasifikasi dengan menerapkan algoritme *k-nearest neighbor*.

3.2.4 Perancangan Algoritme

Perancangan algoritme dilakukan untuk mengidentifikasi beberapa metode yang akan diterapkan dalam penelitian ini. Adapun proses dalam pengimplementasian algoritme pada penelitian ini dimulai dengan menyiapkan

keseluruhan data yang akan dibagi menjadi data latih serta data uji, kemudian dilakukan seleksi fitur untuk mendapatkan fitur terpilih menggunakan *Information Gain* dan selanjutkan akan dilakukan klasifikasi dengan menggunakan algoritme *K-Nearest Neighbor* (K-NN). Perancangan algoritme tersebut ditunjukkan pada Gambar 3.1.



Gambar 3.1 Diagram Alir Proses Klasifikasi

3.2.5 Pengujian Sistem

Pengujian sistem dilakukan setelah penerapan dari metode yang digunakan dalam penelitian. Pengujian bertujuan untuk mengetahui apakah metode yang diterapkan tersebut berhasil atau tidak melalui hasil akurasi dari klasifikasi tersebut. Pengujian dilakukan dengan cara membandingkan antara hasil prediksi dari sistem dengan hasil klasifikasi yang sebenarnya yang nantinya memiliki nilai dan disebut dengan akurasi.

3.2.6 Peralatan Pendukung

Pada sub bab ini akan menguraikan berbagai peralatan pendukung yang digunakan dalam penelitian ini yang mencakup spesifikasi piranti keras (*hardware*) maupun piranti lunak (*software*). Adapun piranti keras yang digunakan diantaranya:

- Intel(R) Pentium(R) CPU 2020M @ 2.40GHz
- RAM 4 GB
- HDD 500 GB
- Monitor 13”

Sedangkan untuk piranti lunak yang digunakan diantaranya:

- Sistem Operasi Windows 8.1
- Microsoft Excel
- Microsoft Word
- Sublime Text

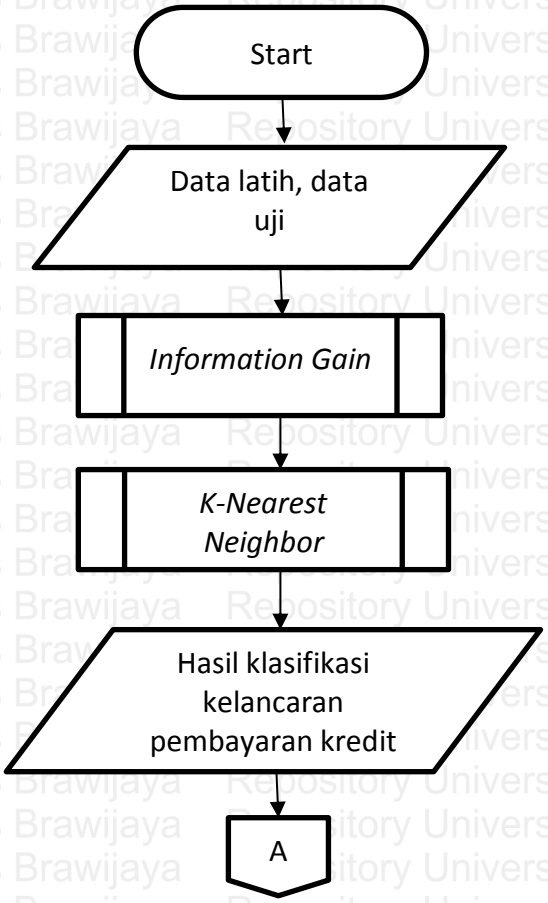
3.2.7 Penarikan Kesimpulan

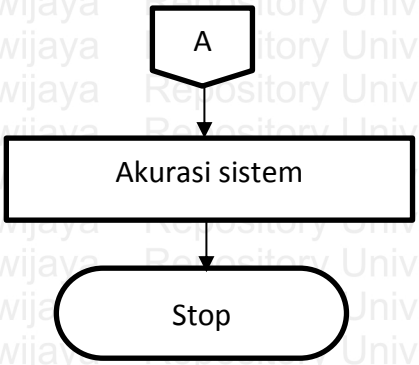
Penarikan kesimpulan dilaksanakan setelah tahap perancangan, tahap implementasi serta tahap pengujian selesai. Kesimpulan memuat jawaban atas permasalahan yang telah dituliskan sebelumnya. Kesimpulan didapatkan mengacu pada hasil pengujian dan analisis dari penerapan metode-metode yang digunakan dalam penelitian ini. Selain kesimpulan, terdapat juga saran dari peneliti sebagai bahan pertimbangan dan perbaikan untuk penelitian berikutnya.

BAB 4 PERANCANGAN

4.1 Perancangan Tahapan Proses

Perancangan dalam klasifikasi tingkat kelancaran pembayaran kredit kendaraan ini adalah dengan menggunakan data kredit kendaraan. Data tersebut akan diproses menggunakan metode seleksi fitur *Information Gain* untuk menghasilkan atribut yang memberikan pengaruh dalam pengklasifikasian tingkat kelancaran pembayaran kredit kendaraan ini sehingga atribut tidak terlalu banyak atribut yang digunakan untuk proses selanjutnya. Setelah mendapatkan atribut-atribut yang ingin digunakan, kemudian dilakukan tahap klasifikasi dengan menggunakan metode *K-Nearest Neighbor* (KNN). Dalam KNN, proses perhitungan dilakukan dengan cara menghitung jarak antara data uji dengan data latih menggunakan perhitungan jarak *euclidean*. Nilai jarak-jarak tersebut kemudian diurutkan mulai dari jarak terkecil atau dengan kata lain data dengan jarak yang paling dekat dan data tersebut diambil sebanyak *K* yang diinginkan. Proses terakhir yaitu memilih kelas mayor atau kelas dengan frekuensi terbanyak yang nantinya dijadikan sebagai penentu dari kelas hasil klasifikasi. Keluaran dari sistem ini berupa klasifikasi tingkat kelancaran pembayaran kredit kendaraan yaitu lancar, diragukan, atau macet. Tahapan dari proses penerapan metode dalam penelitian ini ditunjukkan pada Gambar 4.1.

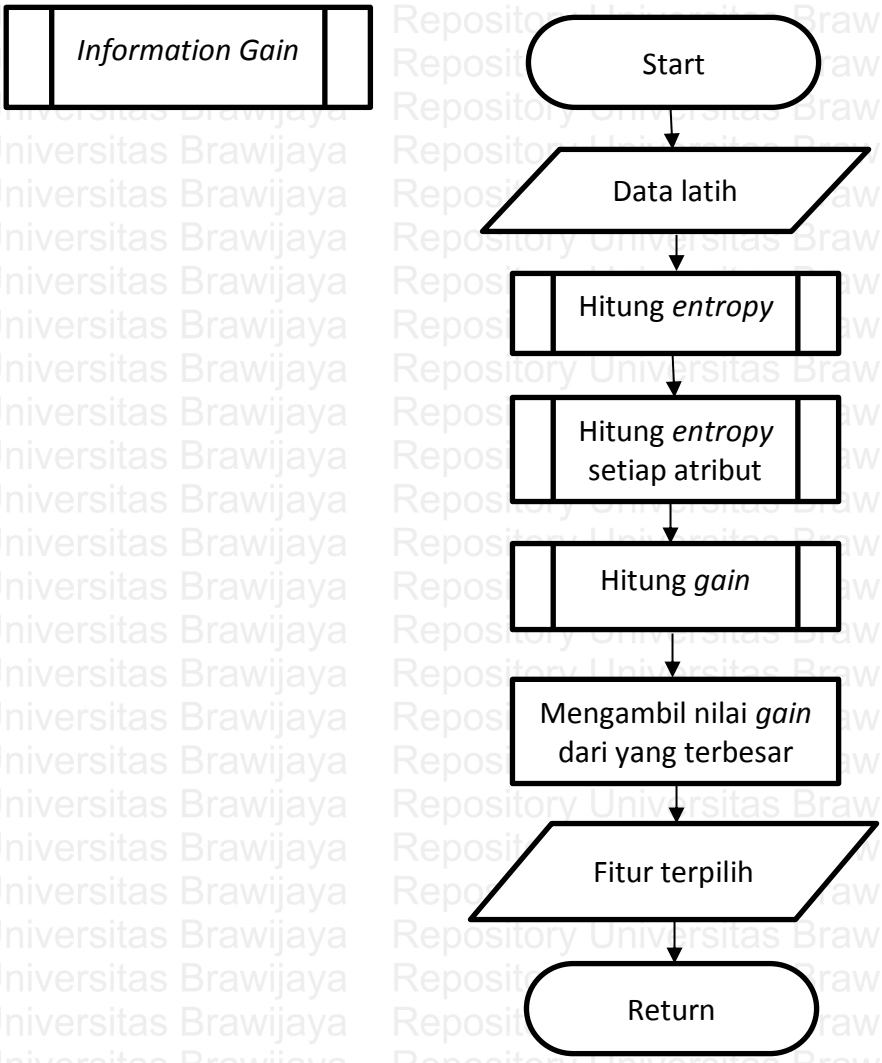




Gambar 4.1 Tahapan Proses Sistem

4.1.1 Tahapan Proses *Information Gain*

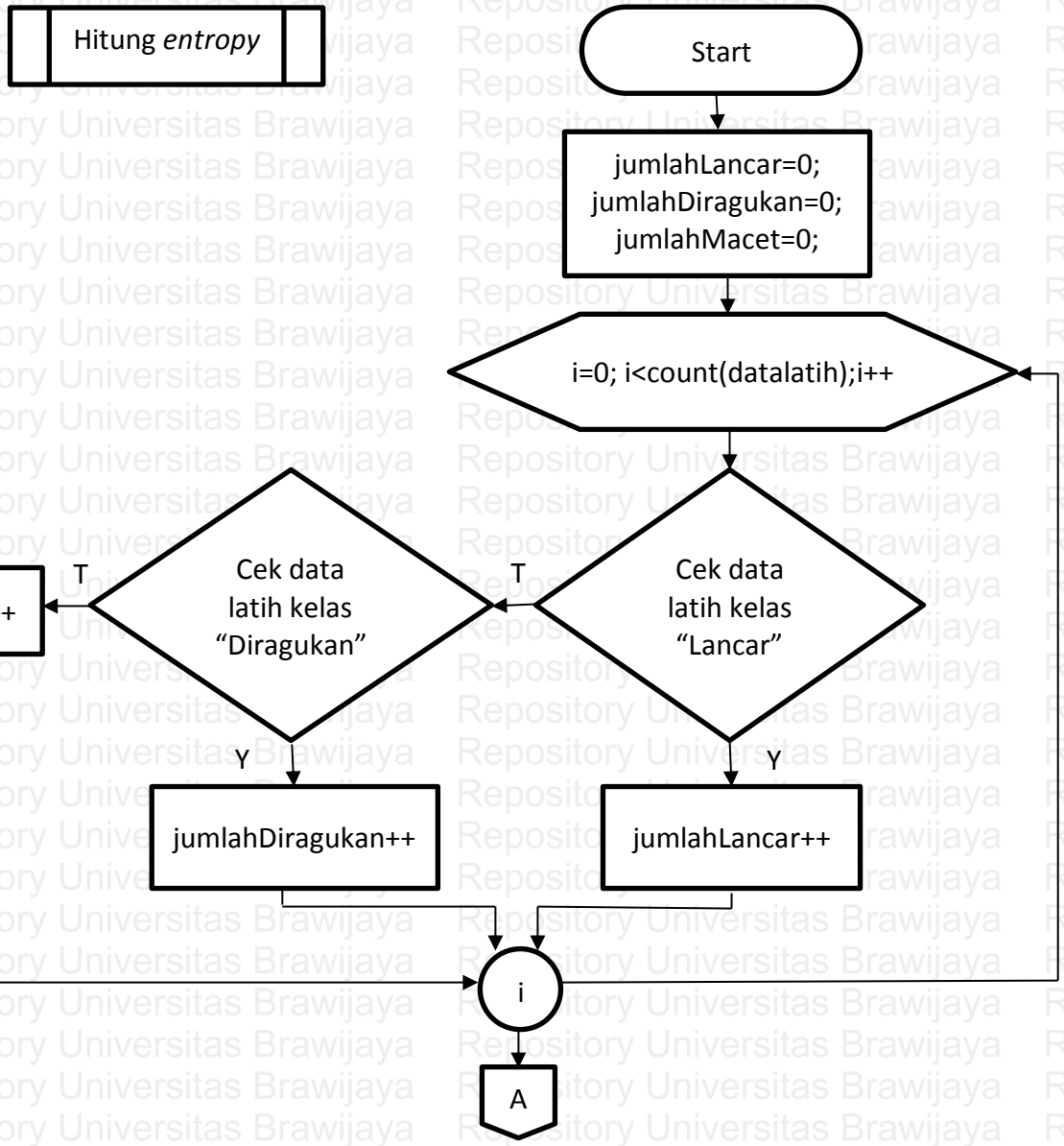
Tahapan proses *Information Gain* diawali dengan melakukan perubahan terhadap data numerik menjadi data kategori. Proses selanjutnya menghitung nilai *entropy* kelas dan setiap atribut kemudian menghitung nilai *gain*. Nilai *gain* tersebut kemudian diurutkan dari nilai terbesar ke terkecil. Gambar 4.2 menunjukkan tahapan proses *Information Gain*.

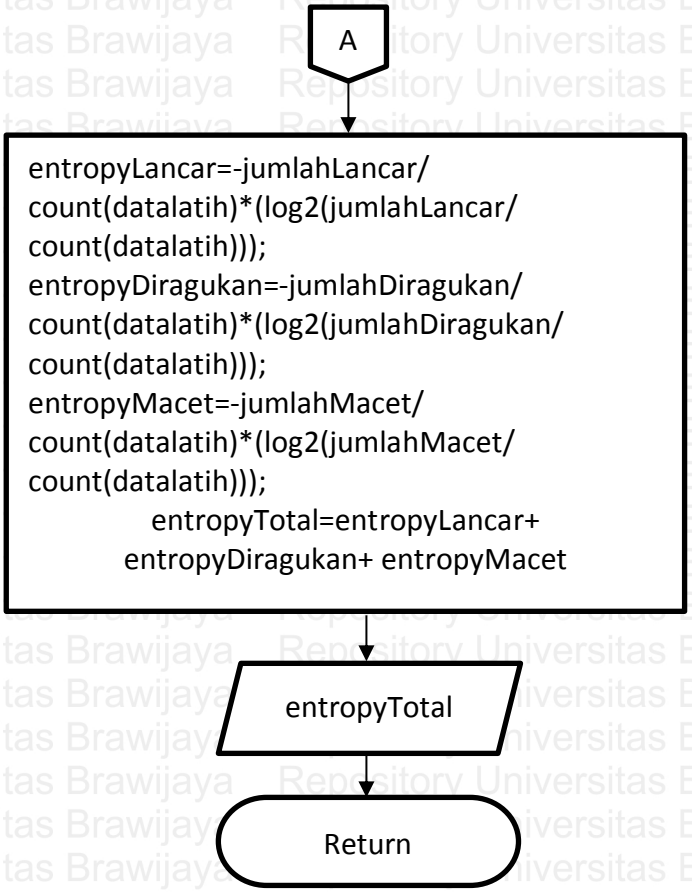


Gambar 4.2 Tahapan Proses *Information Gain*

4.1.1.1 Tahapan Proses Hitung *Entropy*

Tahapan pada proses ini merupakan tahapan dalam proses *Information Gain*. Proses ini dilakukan dengan cara menghitung menggunakan persamaan (2.1) dan akan dijabarkan pada Gambar 4.3. Tahap pertama yaitu mencantumkan data latih, data uji, serta jumlah fitur dan dilakukan inisialisasi pada setiap kelas. Selanjutnya akan dilakukan perulangan dan pengecekan sebanyak jumlah data latih sebagai proses untuk menghitung *entropy* untuk setiap kelas. Proses menghitung *entropy* dilakukan dengan menghitung peluang setiap kelas yang nantinya akan dijumlahkan untuk mendapatkan nilai *entropy* keseluruhan.

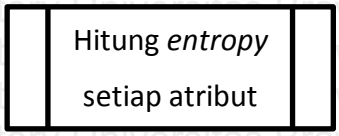


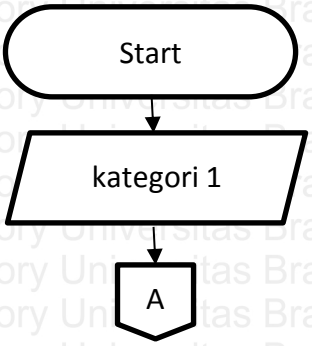


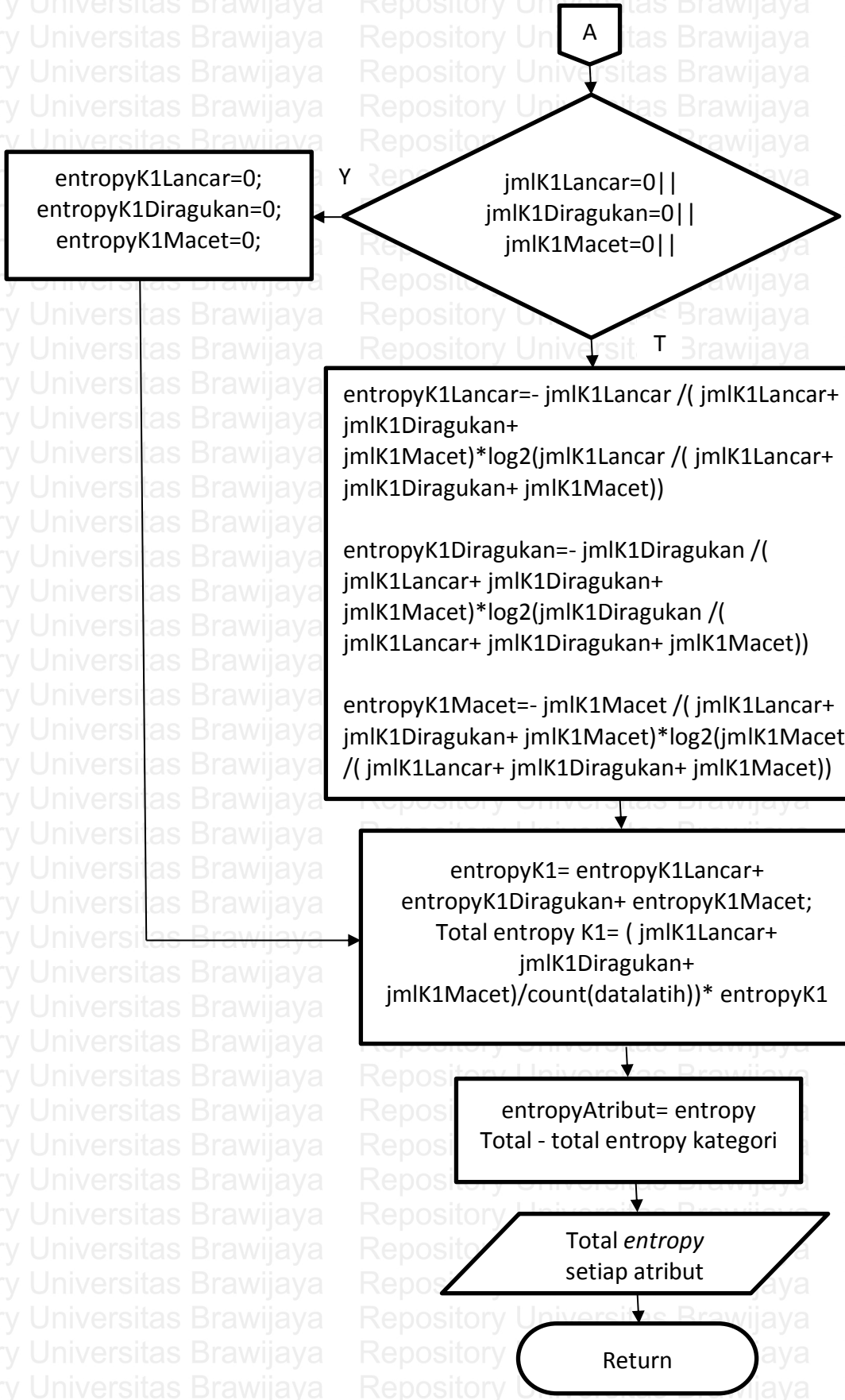
Gambar 4.3 Tahapan Proses Hitung *Entropy*

4.1.1.2 Tahapan Proses Hitung *Entropy* Setiap Atribut

Tahapan proses hitung *entropy* setiap atribut akan dijelaskan pada Gambar 4.4 yang dimulai dengan mencantumkan kategori dari setiap atribut. Selanjutnya dilakukan proses penjumlahan total setiap kategori pada atribut pada setiap kelas. Jika kategori atribut di kelas Lancar, Diragukan, dan Macet bernilai 0, maka peluang setiap atribut dan *entropy* bernilai 0. Sedangkan jika kategori atribut di kelas Lancar, Diragukan, dan Macet tidak bernilai 0, maka akan dilakukan perhitungan peluang setiap kategori atribut pada setiap kelas yang kemudian akan dilakukan perhitungan *entropy* setiap atribut. Misalnya untuk kategori 1 merupakan perempuan pada atribut jenis kelamin. Jika jumlah perempuan pada kelas lancar, diragukan, dan macet tidak ada, maka nilai *entropy* untuk perempuan bernilai 0. Hasil akhir dari tahapan proses ini adalah total *entropy* dari setiap atribut.



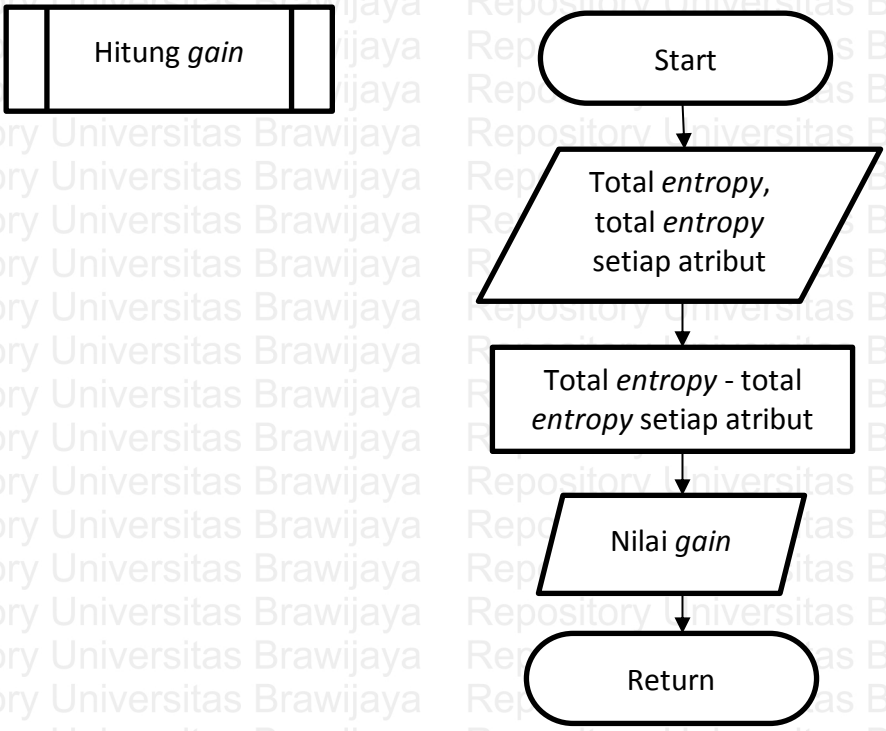




Gambar 4.4 Tahapan Proses Hitung *Entropy* Setiap Atribut

4.1.1.3 Tahapan Proses Hitung *Gain*

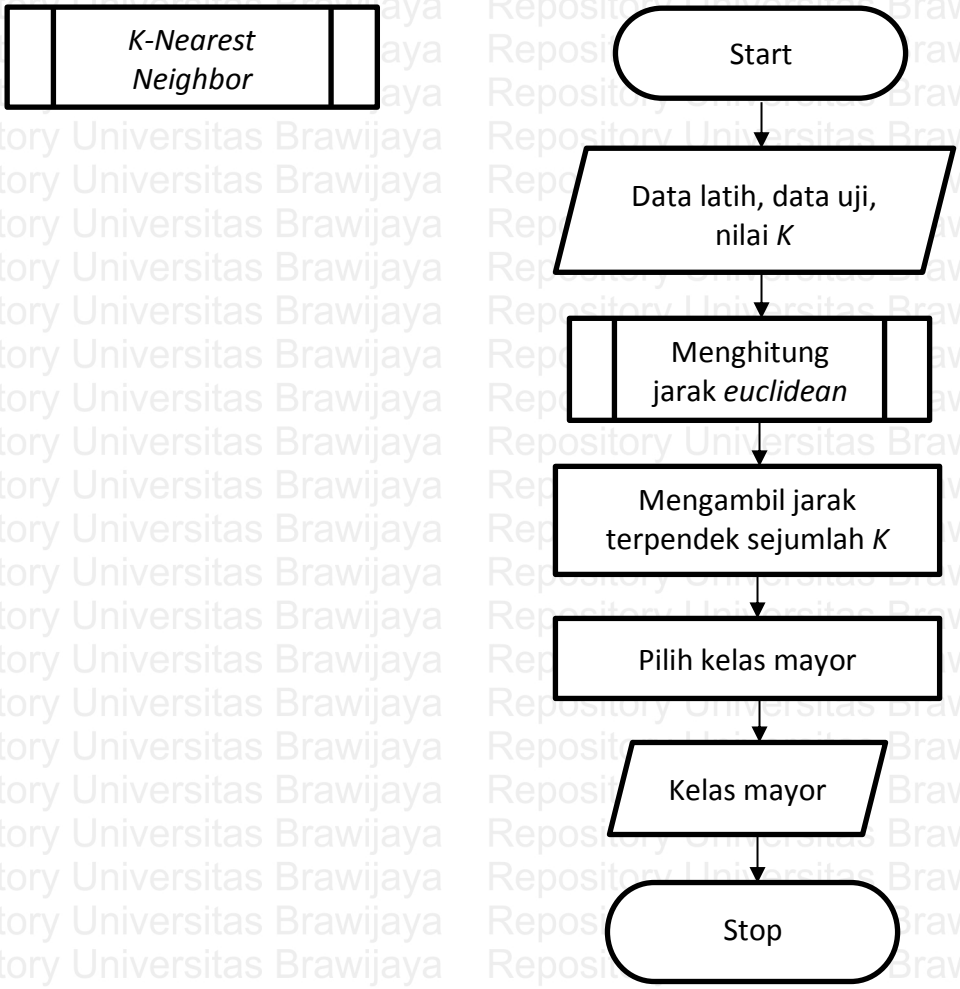
Tahapan proses selanjutnya yaitu menghitung *gain* yang akan ditunjukkan pada Gambar 4.5. Proses diawali dengan mencantumkan nilai *entropy* kelas dan total *entropy* setiap atribut yang selanjutnya dilakukan perhitungan menggunakan persamaan (2.2). Hasil akhir dalam proses ini yaitu nilai gain dari setiap atribut.



Gambar 4.5 Tahapan Proses Hitung *Gain*

4.1.2 Tahapan Proses *K-Nearest Neighbor*

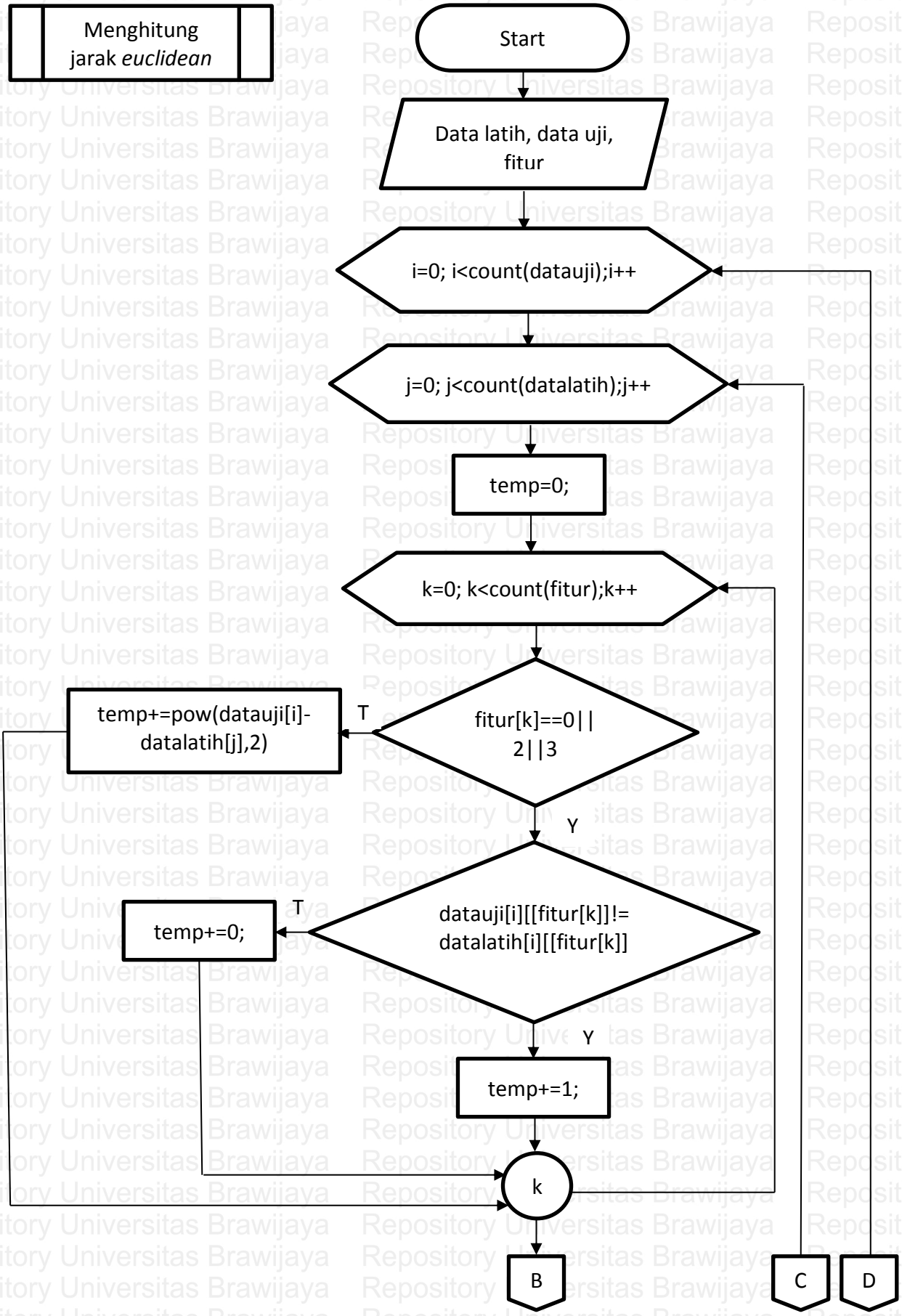
Tahapan proses *K-Nearest Neighbor* dimulai dengan mencantumkan seluruh data yang digunakan dan menentukan nilai *K* (ketetanggan). Data latih dan data uji kemudian dihitung jaraknya menggunakan persamaan *euclidean distance* (2.4). Nilai jarak-jarak yang dihasilkan pada proses sebelumnya kemudian diurutkan berdasarkan nilai jarak terpendek atau terkecil dan diambil sejumlah parameter *K* yang sudah ditetapkan sebelumnya. Kelas yang dominan merupakan hasil dari klasifikasi. Tahapan proses *K-Nearest Neighbor* dapat dilihat pada Gambar 4.7.



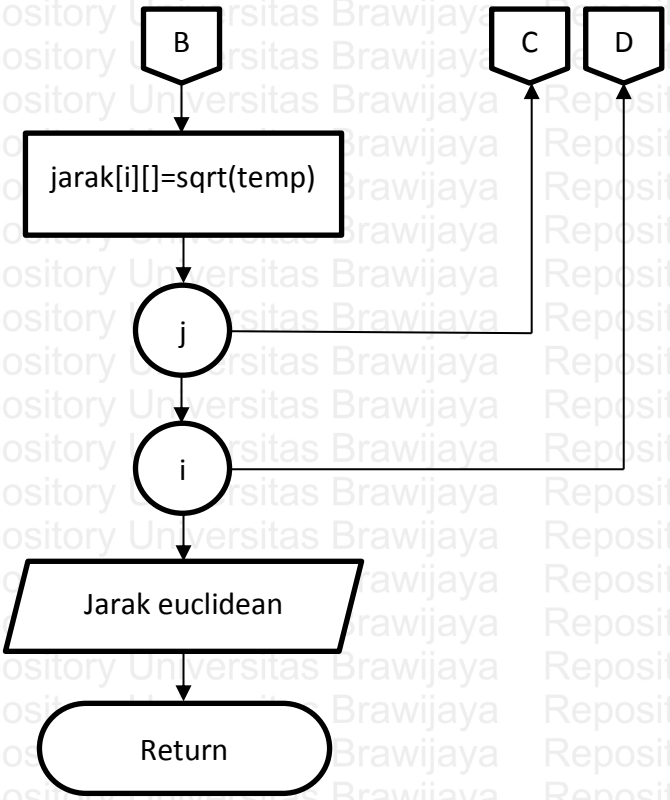
Gambar 4.6 Tahapan Proses *K-Nearest Neighbor*

4.1.2.1 Tahapan Proses Hitung Jarak *Euclidean*

Tahapan proses perhitungan jarak *euclidean* diawali dengan memasukkan data latih dan data uji. Sebelum menghitung jarak *euclidean*, jarak antara data latih dan data uji yang merupakan data kategori didefinisikan menjadi fungsi menggunakan persamaan (2.6). Fungsi tersebut menyatakan jika data latih dan data uji sama berarti jaraknya bernilai 0, sebaliknya jika data latih dan data uji berbeda maka nilai jaraknya bernilai 1. Terdapat tiga fitur yang memiliki data kategori yaitu fitur dengan indeks 0,2, dan 3. Kemudian jarak keseluruhan dihitung dengan menggunakan persamaan (2.5). Hasil akhir dari proses ini merupakan nilai jarak antara data uji dengan data latih. Tahapan proses menghitung jarak *euclidean* dapat dilihat pada Gambar 4.8.



Repository Universitas Brawijaya
REPOSITORY.UB.AC.ID
UNIVERSITAS BRAWIJAYA



Gambar 4.7 Tahapan Proses Hitung Jarak *Euclidean*

4.2 Perhitungan Manual

Perhitungan manual atau sering disebut dengan istilah manualisasi dalam penelitian ini menggunakan 13 data yang dipilih secara acak dari total data yang berjumlah 240. Data tersebut terbagi menjadi 12 data latih dan 1 sisanya menjadi data uji. Data latih terdiri dari 3 label kelas yaitu lancar, diragukan, dan macet. Atribut yang digunakan berjumlah 10, yaitu jenis kelamin, usia, status, pekerjaan, pendapatan, tanggungan, OTR, DP, tenor, dan angsuran. Data latih dan data uji tersebut dapat dilihat pada Tabel 4.1 dan 4.2.

Tabel 4.1 Data Latih

No	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	Klasifikasi
1	Laki-laki	21	Belum Menikah	Wiraswasta	4000000	0	28308600	3300000	29	1338000	Lancar
2	Laki-laki	36	Menikah	Wiraswasta	5000000	2	18508200	1900000	27	896000	Lancar
3	Perempuan	30	Menikah	PNS	3500000	1	17686300	4000000	23	920000	Lancar
4	Perempuan	28	Belum Menikah	Karyawan Swasta	3800000	0	18056000	3100000	29	831000	Lancar
5	Perempuan	36	Widow	Wiraswasta	5900000	0	16551000	1700000	30	750000	Diragukan
6	Perempuan	53	Widow	Tukang Jahit	4000000	2	19629000	2000000	35	863000	Diragukan
7	Laki-laki	32	Belum Menikah	Honorer	4500000	0	17921000	2800000	23	1000000	Diragukan
8	Laki-laki	37	Menikah	PNS	7000000	3	16450000	4000000	14	1188200	Diragukan
9	Perempuan	52	Menikah	Buruh	3300000	6	14259000	1125000	23	788000	Macet
10	Laki-laki	41	Menikah	Wiraswasta	4500000	3	18471000	1850000	29	938000	Macet
11	Perempuan	32	Menikah	Wiraswasta	4500000	5	17574200	2650000	29	820100	Macet
12	Laki-laki	35	Menikah	Wiraswasta	4500000	4	16019000	3250000	29	705600	Macet

Tabel 4.2 Data Uji

No	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	Klasifikasi
96	Perempuan	53	Menikah	PNS	3000000	0	17201000	1800000	11	1816000	?

Repository Universitas Brawijaya
REPOSITORY.UB.AC.ID
UNIVERSITAS BRAWIJAYA

Keterangan untuk inisial pada Tabel 4.1 dan Tabel 4.2 adalah sebagai berikut:

A1= Jenis kelamin
A2= Usia
A3= Status
A4= Pekerjaan
A5= Pendapatan
A6= Tanggungan
A7= Harga On The Road (OTR)
A8= DP
A9= Tenor
A10= Angsuran
Klasifikasi= label kelas kredit kendaraan

Berikut ini langkah-langkah dalam proses perhitungan:

Langkah 1. Konversi data numerik menjadi kategori

Langkah awal yang dilakukan dalam proses perhitungan seleksi fitur *Information Gain* adalah melakukan perubahan dari data numerik pada atribut menjadi data kategori karena data yang digunakan terdiri dari data numerik dan kategori. Adapun atribut numerik yang dikonversi menjadi kategori yaitu:

1. Usia

Atribut usia terbagi menjadi empat tingkatan yaitu remaja, dewasa, lansia, dan manula. Pembagian kategori tersebut diperoleh dari Depkes RI (2009) dan dapat dilihat pada Tabel 4.3.

Tabel 4.3. Konversi Nilai Usia

Usia	Kategori
12 – 25 tahun	Remaja
26 – 45 tahun	Dewasa
46 – 65 tahun	Lansia
> 65 tahun	Manula

2. Pendapatan

Menurut Badan Pusat Statistik (2014), penggolongan pendapatan dapat dibagi kedalam 4 golongan yaitu dangat tinggi, tinggi, sedang, dan rendah. Tabel 4.4 menunjukkan penggolongan dari pendapatan.

Tabel 4.4 Konversi Nilai Pendapatan

Pendapatan	Kategori
> Rp. 3.500.000,00	Sangat tinggi
Rp. 2.500.000,00 – Rp. 3.500.000,00	Tinggi
Rp. 1.500.000,00 – Rp. 2.500.000,00	Sedang
< Rp. 1.500.000,00	Rendah

3. Tanggungan

Atribut tanggungan terbagi menjadi 3 kategori yaitu keluarga kecil, keluarga sedang, dan keluarga besar. Ketiga kategori tersebut diperoleh dari Badan Pusat Statistik. Pembagian kategori untuk atribut tanggungan dapat dilihat pada Tabel 4.5.

Tabel 4.5 Konversi Nilai Tanggungan

Jumlah Tanggungan	Kategori
1 – 3	Keluarga kecil
4 – 6	Keluarga sedang
> 6	Keluarga besar

Pada atribut tanggungan di data yang digunakan, terdapat data dengan jumlah tanggungan sejumlah 0. Sehingga dilakukan penambahan satu kategori yaitu belum berkeluarga. Tabel 4.6 menunjukkan pembagian kategori tanggungan yang digunakan dalam penelitian ini.

Tabel 4.6 Konversi Nilai Tanggungan

Jumlah Tanggungan	Kategori
0	Belum berkeluarga
1 – 3	Keluarga kecil
4 – 6	Keluarga sedang
> 6	Keluarga besar

4. Harga On The Road

Harga On The Road (OTR) merupakan istilah yang digunakan dalam kredit kendaraan yaitu biaya beli suatu kendaraan yang telah ditambahkan dengan biaya pajak dan dokumen lainnya seperti STNK dan BPKB. Atribut OTR dapat dikonversikan menjadi sembilan kategori dengan pembagian kategori berdasarkan konsep perhitungan statistika yaitu mengubah data tunggal menjadi data berkelompok. Proses perhitungan diawali dengan menentukan nilai minimum dan maksimum dari nilai-nilai pada atribut. Kemudian, menentukan jangkauan (R) dengan menghitung selisih antara nilai maksimum dan minimum. Langkah selanjutnya adalah menentukan banyak kelas dengan menggunakan rumus *Sturgess* sebagai berikut:

$$K = 1 + 3{,}3 \log n$$

Dimana:
K : jumlah banyaknya kelas
n : jumlah banyaknya data yang digunakan

Contoh penentuan banyaknya kelas atau kategori pada atribut OTR:
max : 59.951.900
min : 13.204.000
$R = max - min = 59.951.900 - 13.204.000$
$= 46.747.900$
$K = 1 + 3{,}3 \log 240 = 1 + 3{,}3(2{,}380211242)$
$= 8{,}854697098$ dibulatkan menjadi 9 kelas

Langkah terakhir adalah menentukan rentang nilai untuk masing-masing kelas dengan cara membagi nilai jangkauan (R) dengan banyaknya kelas(K). Rentang masing-masing kelas untuk konversi atribut OTR ditunjukkan pada Tabel 4.7.

Tabel 4.7 Konversi Nilai OTR

Harga OTR	Kategori
Rp. 13.204.000 – Rp. 18.398.211	A
Rp. 18.398.212 – Rp. 23.592.422	B
Rp. 23.592.423 – Rp. 28.786.633	C
Rp. 28.786.634 – Rp. 33.980.844	D
Rp. 33.980.845 – Rp. 39.175.056	E
Rp. 39.175.057 – Rp. 44.369.267	F
Rp. 44.369.268 – Rp. 49.563.478	G
Rp. 49.563.478 – Rp. 54.757.689	H
Rp. 54.757.689 – Rp. 59.951.900	I

5. DP

Down Payment atau yang disingkat dengan istilah DP merupakan biaya yang wajib dibayarkan ketika akan mengambil kredit kendaraan. Atribut DP dikonversi menjadi 9 kategori yaitu x1-x9 dengan menggunakan konsep perhitungan statistika yaitu mengubah data tunggal menjadi data berkelompok. Langkah-langkah dalam menentukan rentang nilai untuk masing-masing kelas pada atribut DP sama dengan langkah-langkah yang dilakukan untuk menentukan kategori pada Tabel 4.7. Tabel 4.8 menunjukkan pembagian kategori untuk atribut DP.

Tabel 4.8 Konversi Nilai DP

DP	Kategori
Rp. 725.000 – Rp. 2.611.111	x1
Rp. 2.611.112 – Rp. 4.497.222	x2
Rp. 4.497.223 – Rp. 6.383.333	x3
Rp. 6.383.334 – Rp. 8.269.444	x4
Rp. 8.269.445 – Rp. 10.155.556	x5
Rp. 10.155.557 – Rp. 12.041.667	x6
Rp. 12.041.668 – Rp. 13.927.778	x7
Rp. 13.927.779 – Rp. 15.813.889	x8
Rp. 15.813.890 – Rp. 17.700.000	x9

6. Tenor

Tenor merupakan jangka waktu yang diberikan untuk pelunasan biaya peminjaman kredit yang biasa ditentukan dalam hitungan bulan atau tahun. Atribut tenor dibagi kedalam tiga kategori yang terdiri atas jangka pendek untuk 0-1 tahun, jangka menengah untuk 1-3 tahun dan jangka panjang untuk lebih dari 3 tahun. Tabel 4.9 menunjukkan pembagian kategori atribut tenor.

Tabel 4.9 Konversi Nilai Tenor

Tenor	Kategori
0 – 1 tahun	Jangka pendek
1 – 3 tahun	Jangka menengah
> 3 tahun	Jangka panjang

7. Angsuran

Atribut angsuran dapat dikonversi menjadi 9 kategori yaitu y1-y9. Penentuan pembagian kategori nilai angsuran juga menggunakan konsep perhitungan statistika yaitu mengubah data tunggal menjadi data berkelompok. Langkah-langkah dalam menentukan rentang nilai untuk masing-masing kelas pada atribut angsuran sama dengan langkah-langkah yang dilakukan untuk menentukan kategori pada Tabel 4.8. Tabel 4.10 menunjukkan konversi nilai angsuran.

Tabel 4.10 Konversi Nilai Angsuran

Angsuran	Kategori
Rp. 308.600 – Rp. 585.867	y1
Rp. 585.868 – Rp. 863.133	y2
Rp. 863.134 – Rp. 1.140.400	y3
Rp. 1.140.401 – Rp. 1.417.667	y4
Rp. 1.417.668 – Rp. 1.694.933	y5
Rp. 1.694.934 – Rp. 1.972.200	y6
Rp. 1.972.201 – Rp. 2.249.467	y7
Rp. 2.249.468 – Rp. 2.526.733	y8
Rp. 2.526.734 – Rp. 2.804.000	y9

Langkah 2. Menghitung *entropy*

Menghitung *entropy* dilakukan setelah semua atribut yang berjenis numerik dikonvesi menjadi kategori. Dalam perhitungan *entropy* digunakan persamaan (2.2). Perhitungan *entropy* dilakukan pada kelas data dan pada setiap atribut. Dalam proses ini akan dilakukan perhitungan manual *entropy* untuk kelas Lancar (L), Diragukan (D), dan Macet (M).

Contoh perhitungan *entropy*:

$$Entropy(S) = (-P_L log_2 P_L) + (-P_D log_2 P_D) + (-P_M log_2 P_M)$$
$$= (-\frac{4}{12} log_2 \frac{4}{12}) + (-\frac{4}{12} log_2 \frac{4}{12}) + (-\frac{4}{12} log_2 \frac{4}{12})$$
$$= 0{,}528320834 + 0{,}528320834 + 0{,}528320834$$
$$= 1{,}584962501$$

Langkah 3. Menghitung nilai *Information Gain* setiap atribut

langkah selanjutnya setelah mendapatkan nilai *entropy* kelas yaitu menghitung *entropy* untuk setiap atribut. Setelah nilai entropy dari setiap atribut didapatkan, maka dilakukan perhitungan untuk mendapatkan nilai *gain* dengan

menggunakan persamaan (2.3) dan (2.4). Dalam perhitungan manual ini akan dijabarkan perhitungan untuk mendapatkan *entropy* dan nilai *gain* atribut usia.

Contoh perhitungan *entropy* atribut usia:

a. Kategori Remaja

$$Entropy(S_{Remaja}) = (-P_L log_2 P_L) + (-P_D log_2 P_D) + (-P_M log_2 P_M)$$
$$= (-\tfrac{1}{1} log_2 \tfrac{1}{1}) + (-\tfrac{0}{1} log_2 \tfrac{0}{1}) + (-\tfrac{0}{1} log_2 \tfrac{0}{1})$$
$$= 0 + 0 + 0$$
$$= 0$$

b. Kategori Dewasa

$$Entropy(S_{Dewasa}) = (-P_L log_2 P_L) + (-P_D log_2 P_D) + (-P_M log_2 P_M)$$
$$= (-\tfrac{3}{9} log_2 \tfrac{3}{9}) + (-\tfrac{3}{9} log_2 \tfrac{3}{9}) + (-\tfrac{3}{9} log_2 \tfrac{3}{9})$$
$$= 0{,}528320834 + 0{,}528320834 + 0{,}528320834$$
$$= 1{,}584962501$$

c. Kategori Lansia

$$Entropy(S_{Lansia}) = (-P_L log_2 P_L) + (-P_D log_2 P_D) + (-P_M log_2 P_M)$$
$$= (-\tfrac{0}{0} log_2 \tfrac{0}{0}) + (-\tfrac{1}{2} log_2 \tfrac{1}{2}) + (-\tfrac{1}{2} log_2 \tfrac{1}{2})$$
$$= 0 + 0{,}5 + 0{,}5$$
$$= 1$$

d. Kategori Manula

$$Entropy(S_{Lansia}) = (-P_L log_2 P_L) + (-P_D log_2 P_D) + (-P_M log_2 P_M)$$
$$= (-\tfrac{0}{0} log_2 \tfrac{0}{0}) + (-\tfrac{0}{0} log_2 \tfrac{0}{0}) + (-\tfrac{0}{0} log_2 \tfrac{0}{0})$$
$$= 0 + 0 + 0$$
$$= 0$$

Contoh perhitungan *Information Gain* atribut usia:

$$Gain(S, Usia) = 1{,}584962501 - \{((1 + 0 + 0)/12\ x0) + ((3 + 3 + 3)/12\ x1{,}584962501) + ((0 + 1 + 1)/12\ x1) + ((0 + 0 + 0)/12\ x0)\}$$
$$= 1{,}584962501 - (0 + 1{,}188721876 + 0{,}166666667 + 0)$$
$$= 1{,}584962501 - 1{,}355388542$$
$$= 0{,}229573959$$

Langkah 4. Mengurutkan nilai *Information Gain*

Langkah keempat yaitu mengurutkan atribut-atribut yang berdasarkan nilai *gain* terbesar. Urutan atribut tersebut menetapkan atribut apa saja yang akan digunakan dalam proses klasifikasi. Tabel 4.11 menunjukkan hasil nilai *gain* dari setiap atribut yang telah diurutkan.

Tabel 4.11 Urutan Nilai *Information Gain*

No.	Atribut	Nilai *Gain*
1	Pekerjaan	0,688721875
2	Tanggungan	0,617492461
3	Status	0,551098171
4	Usia	0,229573959
5	Angsuran	0,188721876

No.	Atribut	Nilai Gain
6	OTR	0,147869792
7	Pendapatan	0,109170339
8	DP	0,042776048
9	Jenis Kelamin	0
10	Tenor	0

Langkah 5. Menentukan jumlah atribut

Setelah mengurutkan nilai *gain* dari setiap atribut, maka tahap selanjutnya adalah menetapkan jumlah atribut yang akan digunakan dalam proses klasifikasi. Jumlah atribut yang digunakan dalam perhitungan manual yaitu tiga atribut,maka dari itu atribut-atribut yang terpilih adalah pekerjaan, tanggungan, dan status dikarenakan ketiga atribut tersebut memiliki urutan nilai *gain* teratas.

Langkah 6. Menentukan nilai *K*

Langkah selanjutnya pada proses klasifikasi yaitu menetapkan nilai *K*. Nilai *K* adalah nilai untuk menentukan jumlah ketetanggan. Jumlah tetangga yang digunakan dalam perhitungan manual ini adalah *K*=4. Pemilihan nilai *K* tersebut dikarenakan data penelitian memiliki tiga label kelas sehingga dipilih nilai *K* genap untuk mendapatkan kelas mayoritas.

Langkah 7. Melakukan normalisasi data

Sebelum melakukan perhitungan jarak, data numerik yang ada pada fitur atau atribut dilakukan normalisasi terlebih dahulu. Normalisasi data dilakukan karena data numerik pada penelitian ini memiliki rentang yang jauh pada masing-masing atribut. Normalisasi yang digunakan dalam penelitian ini adalah *min-max normalization*. Data yang akan dilakukan normalisasi ditunjukkan pada Tabel 4.12 dan 4.13.

Tabel 4.12 Data Latih Baru

No.	Status	Pekerjaan	Tanggungan	Klasifikasi
1	Belum Menikah	Wiraswasta	0	Lancar
2	Menikah	Wiraswasta	0	Lancar
3	Menikah	PNS	1	Lancar
4	Belum Menikah	Karyawan Swasta	0	Lancar
5	Widow	Wiraswasta	0	Diragukan
6	Widow	Tukang Jahit	2	Diragukan
7	Belum Menikah	Honorer	0	Diragukan
8	Menikah	PNS	3	Diragukan
9	Menikah	Buruh	6	Macet
10	Menikah	Wiraswasta	3	Macet
11	Menikah	Wiraswasta	5	Macet
12	Menikah	Wiraswasta	4	Macet

Tabel 4.13 Data Uji Baru

No.	Status	Pekerjaan	Tanggungan	Klasifikasi
96	Menikah	PNS	0	?

Contoh perhitungan *min-max normalization* atribut tanggungan untuk data latih pertama:

$X = 0$
$X_{min} = 0$
$X_{max} = 6$

$$X' = \frac{0-0}{6-0} = \frac{0}{6} = 0$$

Tabel 4.14 dan 4.15 menunjukkan data latih dan data baru yang telah dinormalisasi.

Tabel 4.14 Data Latih Baru Setelah Normalisasi

No.	Status	Pekerjaan	Tanggungan	Klasifikasi
1	Belum Menikah	Wiraswasta	0	Lancar
2	Menikah	Wiraswasta	0,333333333	Lancar
3	Menikah	PNS	0,166666667	Lancar
4	Belum Menikah	Karyawan Swasta	0	Lancar
5	Widow	Wiraswasta	0	Diragukan
6	Widow	Tukang Jahit	0,333333333	Diragukan
7	Belum Menikah	Honorer	0	Diragukan
8	Menikah	PNS	0,5	Diragukan
9	Menikah	Buruh	1	Macet
10	Menikah	Wiraswasta	0,5	Macet
11	Menikah	Wiraswasta	0,833333333	Macet
12	Menikah	Wiraswasta	0,666666667	Macet

Tabel 4.15 Data Uji Baru Setelah Normalisasi

No.	Status	Pekerjaan	Tanggungan	Klasifikasi
96	Menikah	PNS	0	?

Langkah 8. Menghitung jarak antara data uji dengan data latih

Langkah berikutnya masih dalam tahap klasifikasi yaitu menghitung jarak kedekatan antara data uji baru dengan data latih baru. Proses menghitung jarak pada penelitian ini menggunakan persamaan jarak *eulidean*(2.5).

Contoh perhitungan jarak antara data latih pertama ke data uji:

$$D(1{,}96) = \sqrt{1^2 + 1^2 + (0-0)^2} = 1{,}414213562$$

Tabel 4.16 menunjukkan hasil dari perhitungan jarak kedekatan antara semua data latih dengan data uji.

Tabel 4.16 Hasil Perhitungan Jarak

	Jarak	Klasifikasi
d(1,96)	1,414213562	Lancar
d(2,96)	1,054092553	Lancar
d(3,96)	0,166666667	Lancar
d(4,96)	1,414213562	Lancar
d(5,96)	1,414213562	Diragukan
d(6,96)	1,452966315	Diragukan
d(7,96)	1,414213562	Diragukan
d(8,96)	0,5	Diragukan
d(9,96)	1,414213562	Macet
d(10,96)	1,118033989	Macet
d(11,96)	1,301708279	Macet
d(12,96)	1,201850425	Macet

Langkah 8. Mengurutkan jarak

Setelah mendapatkan nilai jarak semua data latih dengan data uji, langkah selanjutnya adalah mengurutkan jarak data dari yang terkecil yang berarti antara dua data sangat dekat. Hasil pengurutan jarak tersebut dapat dilihat pada Tabel 4.17.

Tabel 4.17 Hasil Pengurutan Jarak

	Jarak	Klasifikasi
d(3,96)	0,166666667	Lancar
d(8,96)	0,5	Diragukan
d(2,96)	1,054092553	Lancar
d(10,96)	1,118033989	Macet
d(12,96)	1,201850425	Macet
d(11,96)	1,301708279	Macet
d(1,96)	1,414213562	Lancar
d(4,96)	1,414213562	Lancar
d(5,96)	1,414213562	Diragukan
d(7,96)	1,414213562	Diragukan
d(9,96)	1,414213562	Macet
d(6,96)	1,452966315	Diragukan

Langkah 9. Mengambil data sejumlah *K*

Setelah didapatkan nilai jarak yang terurut, tahap selanjutnya dari proses *K-Nearest Neighbor* yaitu mengambil data sejumlah nilai *K* yang telah ditentukan pada langkah 6. Pada langkah 6, tetanggaan yang ditentukan sejumlah 4 yang

berarti sejumlah 4 data yang memiliki jarak terpendek. Hasil pemilihan data sebanyak *K* dapat dilihat pada Tabel 4.18.

Tabel 4.18 Pemilihan Data Sebanyak *K*

	Jarak	Klasifikasi
d(3,96)	0,166666667	Lancar
d(8,96)	0,5	Diragukan
d(2,96)	1,054092553	Lancar
d(10,96)	1,118033989	Macet

Langkah 10. Pemilihan Kelas Mayor

Proses klasifikasi ditutup dengan proses terakhir yaitu penentuan kelas untuk data uji berdasarkan frekuensi kelas terbanyak pada langkah sebelumnya atau biasa disebut kelas mayor. Berdasarkan Tabel 4.18 kelas mayor yang diperoleh adalah Lancar dikarenakan terdapat dua nilai jarak dengan hasil klasifikasi Lancar sehingga data uji tersebut tergolong ke dalam label kelas Lancar. Hasil akhir proses klasifikasi dapat dilihat pada Tabel 4.19.

Tabel 4.19 Hasil Manualisasi

No.	Status	Pekerjaan	Tanggungan	Klasifikasi
96	Menikah	PNS	0	Lancar

4.3 Perancangan Pengujian

Pengujian dilakukan untuk mengukur tingkat akurasi dari sistem yang telah dibuat. Pengujian akurasi dilakukan berdasarkan jumlah fitur dan nilai *K* yang diinputkan. Tabel 4.20 menunjukkan perancangan pengujian.

Tabel 4.20 Perancangan Pengujian

Jumlah Fitur	Akurasi (%)				
	K=	*K*=	*K*=	...	*K*=
...					

Repository Universitas Brawijaya
REPOSITORY.UB.AC.ID
UNIVERSITAS BRAWIJAYA

BAB 5 IMPLEMENTASI

5.1 Spesifikasi Sistem

Pada sub bab ini akan menjelaskan mengenai rincian dari perangkat yang digunakan dalam pengimplemntasian metode yang digunakan dalam penelitian. Spesifikasi perangkat terdiri atas dua, yaitu perangkat keras dan perangkat lunak.

5.1.1 Spesifikasi Perangkat Keras

Spesifikasi perangkat keras yang digunakan dalam sistem klasifikasi tingkat kelancaran pembayaran kredit kendaraan ini ditunjukkan pada Tabel 5.1.

Tabel 5.1 Spesifikasi Perangkat Keras

Nama Komponen	Spesifikasi
Prosesor	Intel(R) Pentium(R) CPU 2020M @ 2.40 GHz
RAM	4.00 GB
Hardisk	HDD 500GB

5.1.2 Spesifikasi Perangkat Lunak

Spesifikasi perangkat keras yang digunakan dalam sistem klasifikasi tingkat kelancaran pembayaran kredit kendaraan ini ditunjukkan pada Tabel 5.2.

Tabel 5.2 Spesifikasi Perangkat Lunak

Nama	Spesifikasi
Sistem Operasi	Windows 8.1
Bahasa Pemrograman	PHP
Tools Pemrograman	Sublime Text 3

5.2 Batasan Implementasi

Batasan-batasan dalam implementasi metode dalam penelitian ini adalah sebagai berikut:

1. Bahasa pemrograman yang digunakan dalam implementasi penelitian adalah PHP.
2. Metode dalam penelitian ini menggunakan dua metode. *Information Gain* sebagai metode untuk seleksi fitur dan *K-Nearest Neighbor* sebagai metode untuk klasifikasi.
3. Masukan dalam sistem terdiri dari dua yaitu jumlah fitur yang akan digunakan untuk proses klasifikasi berdasarkan hasil dari perangkingan perhitungan *Information Gain* dan nilai *K*.
4. Keluaran dari sistem berupa label kelas klasifikasi tingkat kelancaran pembayaran kredit kendaraan yang terdiri dari tiga kelas, yaitu Lancar, Diragukan, dan Macet.

Repository Universitas Brawijaya
REPOSITORY.UB.AC.ID
UNIVERSITAS BRAWIJAYA

5.3 Implementasi Algoritme

Pada sub bab ini menjelaskan mengenai proses penerapan algoritme ke dalam bahasa pemrograman bergantung pada proses perancangan algoritme yang sebelumnya telah diuraikan pada bab 4. Dalam penelitian ini, terdapat 2 metode yaitu *Information Gain* untuk seleksi fitur dan *K-Nearest Neighbor* untuk klasifikasi.

5.3.1 Implementasi Algoritme *Information Gain*

Dalam pengimplementasian metode *Information Gain* yang digunakan untuk mereduksi fitur terdapat dua langkah untuk mendapatkan nilai *gain* dari atribut atau fitur. Langkah pertama adalah perhitungan *entropy* pada kelas data dan langkah kedua adalah perhitungan entropy setiap atribut sehingga mendapatkan nilai *gain*.

5.3.1.1 Implementasi Algoritme Perhitungan *Entropy*

Tahap awal dalam mengimplemntasikan metode *Information Gain* adalah menghitung *entropy* dari setiap kelas pada data yang digunakan. Menghitung *entropy* dari setiap kelas dilakukan dengan cara menghitung peluang munculnya setiap kelas dan dikalikan dengan log^2 dari peluang munculnya setiap kelas tersebut. Dalam penelitian ini terdapat 3 kelas yaitu Lancar, Diragukan, dan Macet. Penjumlahan antara *entropy* Lancar, *entropy* Diragukan, dan *entropy* Macet akan menghasilkan nilai *entropy* keseluruhan. Kode program 5.1 menunjukkan implementasi untuk perhitungan *entropy*.

Langkah 1: Perhitungan *Entropy*

```
1  $jumlah = 0;
2  for ($i = 0; $i < count($datalatih); $i++) {
3      if ($kelas_latih[$i] == "Lancar") {
4          $jumlah = $jumlah + 1;
5      }
6  }
7  $entropyLancar = -$jumlah / count($datalatih) * (log($jumlah /
   count($datalatih), 2));

8  $jumlah = 0;
9  for ($i = 0; $i < count($datalatih); $i++) {
10     if ($kelas_latih[$i] == "Diragukan") {
11         $jumlah = $jumlah + 1;
12     }
13 }
14 $entropyDiragukan = -$jumlah / count($datalatih) * (log($jumlah /
   count($datalatih), 2));

15 $jumlah = 0;
16 for ($i = 0; $i < count($datalatih); $i++) {
17     if ($kelas_latih[$i] == "Macet") {
18         $jumlah = $jumlah + 1;
19     }
20 }
21 $entropyMacet = -$jumlah / count($datalatih) * (log($jumlah /
   count($datalatih), 2));

22 $entropyTotal = $entropyLancar + $entropyDiragukan +
   $entropyMacet;";
```

Kode Program 5.1. Implementasi Algoritme Perhitungan *Entropy*

Penjelasan dari Kode Program 5.1. adalah sebagai berikut:

Baris 1 : proses mengeset nilai awal variabel yang digunakan untuk menghitung peluang kelas Lancar.
Baris 2-6 : proses perhitungan jumlah kelas Lancar.
Baris 7 : proses perhitungan *entropy* untuk kelas Lancar.
Baris 8 : proses mengeset nilai awal variabel yang digunakan untuk menghitung peluang kelas Diragukan.
Baris 9-13 : proses perhitungan jumlah kelas Diragukan.
Baris 14 : proses perhitungan *entropy* untuk kelas Lancar.
Baris 15 : proses mengeset nilai awal variabel yang digunakan untuk menghitung peluang kelas Macet.
Baris 16-20 : proses perhitungan jumlah kelas Macet.
Baris 21 : proses perhitungan *entropy* untuk kelas Macet.
Baris 22 : proses perhitungan total *entropy* dengan menjumlahkan *entropy* Lancar, *entropy* Diragukan, dan *entropy* Macet.

5.3.1.2 Implementasi Algoritme *Entropy* Setiap Atribut

Tahap selanjutnya setelah mendapatkan nilai entropy keseluruhan adalah menghitung nilai *entropy* untuk setiap atribut yang digunakan. Implementasi pada tahap ini kurang lebih sama dengan pengimplementasian tahap sebelumnya. Setiap kategori yang terdapat di dalam atribut dihitung kemunculannya pada setiap kelas. Peluang tersebut kemudian dikalikan dengan log^2 dari peluang itu sendiri dan implementasinya dapat dilihat pada Kode Progran 5.2.

Langkah 2: Perhitungan *Entropy* Setiap Atribut

```
1  $jumlahPL = 0;
2  $jumlahPD = 0;
3  $jumlahPM = 0;
4  $entropyPL = 0;
5  $entropyPD = 0;
6  $entropyPM = 0;
7  for ($i = 0; $i < count($datalatih); $i++) {
       if ($kelas_latih[$i] == "Lancar" && $datalatih[$i][0] ==
   "Perempuan") {
8          $jumlahPL = $jumlahPL + 1;
9      }
10     if ($kelas_latih[$i] == "Diragukan" && $datalatih[$i][0] ==
   "Perempuan") {
11         $jumlahPD = $jumlahPD + 1;
12     }
13     if ($kelas_latih[$i] == "Macet" && $datalatih[$i][0] ==
   "Perempuan") {
14         $jumlahPM = $jumlahPM + 1;
15     }
16 }
17 if ($jumlahPL == 0) {
18     $entropyPL == 0;
19 } else {
20     $entropyPL = -$jumlahPL / ($jumlahPL + $jumlahPD +
   $jumlahPM) * log($jumlahPL / ($jumlahPL + $jumlahPD +
   $jumlahPM), 2);
21 }

22 if ($jumlahPD == 0) {
```

```
23      $entropyPD == 0;
24  } else {
        $entropyPD = -$jumlahPD / ($jumlahPL + $jumlahPD +
25  $jumlahPM) * log($jumlahPD / ($jumlahPL + $jumlahPD +
    $jumlahPM), 2);
    }
26
    if ($jumlahPM == 0) {
27      $entropyPM == 0;
28  } else {
29      $entropyPM = -$jumlahPM / ($jumlahPL + $jumlahPD +
30  $jumlahPM) * log($jumlahPM / ($jumlahPL + $jumlahPD +
    $jumlahPM), 2);
31  }

32  $entropyP = $entropyPL + $entropyPD + $entropyPM;

33  $jumlahLL = 0;
34  $jumlahLD = 0;
35  $jumlahLM = 0;
36  $entropyLL = 0;
37  $entropyLD = 0;
38  $entropyLM = 0;
39  for ($i = 0; $i < count($datalatih); $i++) {
        if ($kelas_latih[$i] == "Lancar" && $datalatih[$i][0] ==
    "Laki-laki") {
40          $jumlahLL = $jumlahLL + 1;
41      }
42      if ($kelas_latih[$i] == "Diragukan" && $datalatih[$i][0] ==
    "Laki-laki") {
43          $jumlahLD = $jumlahLD + 1;
44      }
45      if ($kelas_latih[$i] == "Macet" && $datalatih[$i][0] ==
    "Laki-laki") {
46          $jumlahLM = $jumlahLM + 1;
47      }
48  }

49  if ($jumlahLL == 0) {
50      $entropyLL == 0;
51  } else {
52      $entropyLL = -$jumlahLL / ($jumlahLL + $jumlahLD +
    $jumlahLM) * log($jumlahLL / ($jumlahLL + $jumlahLD +
    $jumlahLM), 2);
53  }

54  if ($jumlahLD == 0) {
55      $entropyLD == 0;
56  } else {
57      $entropyLD = -$jumlahLD / ($jumlahLL + $jumlahLD +
    $jumlahLM) * log($jumlahLD / ($jumlahLL + $jumlahLD +
    $jumlahLM), 2);
58  }

59  if ($jumlahLM == 0) {
60      $entropyLM == 0;
61  } else {
62      $entropyLM = -$jumlahLM / ($jumlahLL + $jumlahLD +
    $jumlahLM) * log($jumlahLM / ($jumlahLL + $jumlahLD +
    $jumlahLM), 2);
63  }

64  $entropyL = $entropyLL + $entropyLD + $entropyLM;
```

Kode Program 5.2. Implementasi Algoritme Perhitungan *Entropy* Setiap Atribut

Penjelasan dari Kode Program 5.2. adalah sebagai berikut:

Baris 1-6 : proses mengeset nilai awal variabel yang digunakan untuk perhitungan *entropy* pada atribut jenis kelamin untuk kategori perempuan.
Baris 7-9 : proses perhitungan jumlah kategori perempuan pada kelas Lancar.
Baris 10-12 : proses perhitungan jumlah kategori perempuan pada kelas Diragukan.
Baris 13-16 : proses perhitungan jumlah kategori perempuan pada kelas Macet.
Baris 17-21 : proses perhitungan *entropy* pada atribut jenis kelamin untuk kategori perempuan pada kelas Lancar.
Baris 22-26 : proses perhitungan *entropy* pada atribut jenis kelamin untuk kategori perempuan pada kelas Diragukan.
Baris 27-31 : proses perhitungan *entropy* pada atribut jenis kelamin untuk kategori perempuan pada kelas Macet.
Baris 32 : proses perhitungan total *entropy* pada atribut jenis kelamin untuk kategori perempuan.
Baris 33-38 : proses mengeset nilai awal variabel yang digunakan untuk perhitungan *entropy* pada atribut jenis kelamin untuk kategori laki-laki.
Baris 39-41 : proses perhitungan jumlah kategori laki-laki pada kelas Lancar.
Baris 42-44 : proses perhitungan jumlah kategori laki-laki pada kelas Diragukan.
Baris 45-48 : proses perhitungan jumlah kategori laki-laki pada kelas Macet.
Baris 49-53 : proses perhitungan *entropy* pada atribut jenis kelamin untuk kategori laki-laki pada kelas Lancar.
Baris 54-58 : proses perhitungan *entropy* pada atribut jenis kelamin untuk kategori laki-laki pada kelas Diragukan.
Baris 59-63 : proses perhitungan *entropy* pada atribut jenis kelamin untuk kategori laki-laki pada kelas Macet.
Baris 64 : proses perhitungan untuk mendapatkan total *entropy* pada atribut jenis kelamin untuk kategori laki-laki.

5.3.1.3 Implementasi Algoritme Perhitungan *Gain*

Tahap terakhir dalam penerapan metode seleksi fitur *Information Gain* yaitu menghitung nilai *gain* pada setiap atribut yang diperoleh dengan cara melakukan perhitungan selisih antara *entropy* keseluruhan dengan *entropy* setiap atribut. Implementasi untuk mendapatkan nilai *gain* setiap atribut dapat dilihat pada Kode Program 5.3.

Langkah 3: Perhitungan *Gain*

```
1  $gainJK = $entropyTotal - ($entropyP + $entropyL);
2  $gainUsia = $entropyTotal - ($entropyR + $entropyD +
   $entropyL + $entropyM);
3
```

```
    $gainStatus = $entropyTotal - ($entropyBN + $entropyN +
    $entropyW);
4   $gainPekerjaan = $entropyTotal - ($entropyPJ + $entropyPT
    + $entropyPD + $entropyH + $entropyPNS + $entropyK +
    $entropyKS + $entropyPP + $entropyWS + $entropyPS +
    $entropyPU + $entropyM + $entropyB + $entropyMS +
    $entropyP + $entropyIRT + $entropyBU + $entropyBA +
    $entropyS);
5   $gainPendapatan = $entropyTotal - ($entropyST + $entropyT
    + $entropyS + $entropyR);
6   $gainTanggungan = $entropyTotal - ($entropyBB + $entropyKK
    + $entropyKS + $entropyKB);
7   $gainOTR = $entropyTotal - ($entropyA + $entropyB +
    $entropyC + $entropyD + $entropyE + $entropyF + $entropyG
    + $entropyH + $entropyI);
8   $gainTenor = $entropyTotal - ($gainJPD + $gainJM +
    $gainJPJ);
9   $gainDP = $entropyTotal - ($entropyx1 + $entropyx2 +
    $entropyx3 + $entropyx4 + $entropyx5 + $entropyx6 +
    $entropyx7 + $entropyx8 + $entropyx9);
10  $gainAngsuran = $entropyTotal - ($entropyy1 + $entropyy2
    + $entropyy3 + $entropyy4 + $entropyy5 + $entropyy6 +
    $entropyy7 + $entropyy8 + $entropyy9);
```

Kode Program 5.3 Implementasi Algoritme Perhitungan Gain

Penjelasan dari Kode Program 5.3. adalah sebagai berikut:

Baris 1 : proses perhitungan *gain* untuk atribut Jenis Kelamin. Perhitungan tersebut diperoleh dari selisih *entropy* keseluruhan dengan total penjumlahan *entropy* kategori pada atribut Jenis Kelamin.

Baris 2 : proses perhitungan *gain* untuk atribut Usia. Perhitungan tersebut diperoleh dari selisih *entropy* keseluruhan dengan total penjumlahan *entropy* kategori pada atribut Usia.

Baris 3 : proses perhitungan *gain* untuk atribut Status. Perhitungan tersebut diperoleh dari selisih *entropy* keseluruhan dengan total penjumlahan *entropy* kategori pada atribut Status.

Baris 4 : proses perhitungan *gain* untuk atribut Pekerjaan. Perhitungan tersebut diperoleh dari selisih *entropy* keseluruhan dengan total penjumlahan *entropy* kategori pada atribut Pekerjaan.

Baris 5 : proses perhitungan *gain* untuk atribut Pendapatan. Perhitungan tersebut diperoleh dari selisih *entropy* keseluruhan dengan total penjumlahan *entropy* kategori pada atribut Pendapatan.

Baris 6 : proses perhitungan *gain* untuk atribut Tanggungan. Perhitungan tersebut diperoleh dari selisih *entropy* keseluruhan dengan total penjumlahan *entropy* kategori pada atribut Tanggungan.

Baris 7 : proses perhitungan *gain* untuk atribut OTR. Perhitungan tersebut diperoleh dari selisih *entropy* keseluruhan dengan total penjumlahan *entropy* kategori pada atribut OTR.

Baris 8 : proses perhitungan *gain* untuk atribut Tenor. Perhitungan tersebut diperoleh dari selisih *entropy* keseluruhan dengan total penjumlahan *entropy* kategori pada atribut Tenor.

Baris 9 : proses perhitungan *gain* untuk atribut DP. Perhitungan tersebut diperoleh dari selisih *entropy* keseluruhan dengan total penjumlahan *entropy* kategori pada atribut DP.

Baris 10 : proses perhitungan *gain* untuk atribut Angsuran. Perhitungan tersebut diperoleh dari selisih *entropy* keseluruhan dengan total penjumlahan *entropy* kategori pada atribut Angsuran.

5.3.2 Implementasi Algoritme *K-Nearest Neighbor*

Dalam pengimplmentasian algoritme *K-Nearest Neighbor* yang digunakan untuk klasifikasi terdapat langkah utama dalam penerapannya yaitu menghitung jarak kedekatan antar data latih dengan data uji.

5.3.2.1 Implementasi Algoritme Perhitungan *Euclidean Distance*

Pada sub ini akan membahas mengenai penerapan dari perancangan tahapan proses *K-Nearest Neighbor* yang telah dijelaskan pada bab sebelumnya. Dalam pengimplementasian metode *K-Nearest Neighbor* terdapat langkah untuk menghitung jarak kedekatan antar data. Metode perhitungan jarak yang diterapkan pada penelitian ini adalah *euclidean distance* yang pengimplementasiannya ditunjukkan pada Kode Program 5.4.

Langkah 1: Perhitungan *Euclidean Distance*

```
1  for ($i = 0; $i < count($datauji); $i++) {
2          for ($j = 0; $j < count($datalatih); $j++) {
3              $temp = 0;
4              for ($k = 0; $k < count($fiturterpilih); $k++) {
5                  if     ($fiturterpilih[$k]     ==     0     ||
   $fiturterpilih[$k] == 2 || $fiturterpilih[$k] == 3) {
                       //echo "kategori";
6                      if   ($datauji[$i][$fiturterpilih[$k]]   !=
   $datalatih[$j][$fiturterpilih[$k]]){
7                          $temp+=1;
                           //echo "selisihnya = 1";
8                      } else {
9                          $temp+=0;
                           //echo "selisihnya = 0";
10                     }
11                 } else {
                       //echo "numerik";
12                     $temp+=pow(($datauji[$i][$fiturterpilih[$k]]
   - $datalatih[$j][$fiturterpilih[$k]]), 2);
13                 }
14             }
15             $jarak[$i][] = sqrt($temp);
16         }
```

Kode Program 5.4 Implementasi Algoritme *Euclidean Distance*

Penjelasan dari Kode Program 5.4. adalah sebagai berikut:

Baris 1-4 : proses perulangan sebanyak jumlah data uji, data latih, dan fitur terpilih.

Baris 5 : proses perhitungan jika fitur terpilih merupakan fitur dengan indeks ke 0,2, atau 3.

Baris 6-10 : proses pendefinisian nilai jarak pada fitur terpilih. Jika nilai data latih sama dengan data uji maka jaraknya bernilai 0, jika sebaliknya maka bernilai 1.

Baris 12-16 : proses perhitungan jarak seluruh fitur dengan menggunakan jarak *euclidean*.

5.4 Implementasi Antarmuka

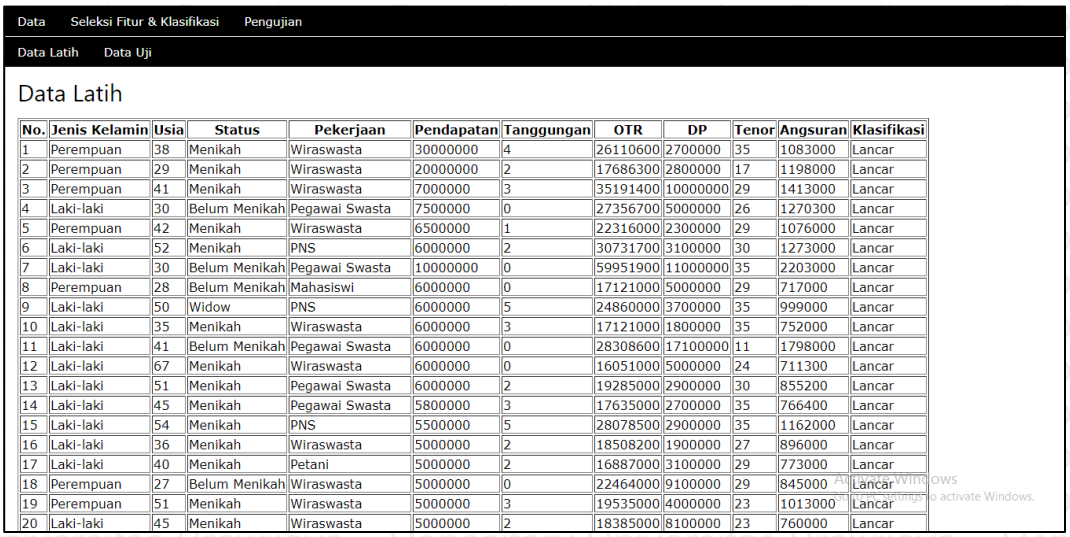
Antarmuka sangat berperan penting bagi pengguna dalam interaksinya dengan sistem. Pada penelitian ini, dibuat antarmuka sederhana yang dibagi menjadi tiga bagian utama. Bagian utama tersebut terdiri dari antarmuka untuk melihat data, antarmuka dalam penerapan metode yang digunakan, dan antarmuka pengujian.

5.4.1 Implementasi Antarmuka Data

Antarmuka data akan menampilkan dua tombol ketika dijalankan yaitu tombol data latih yang akan menampilkan tabel data yang akan digunakan sebagai data latih dan data uji yang akan menampilkan tabel data yang akan digunakan sebagai data uji.

5.4.1.1 Implementasi Antarmuka Data Latih

Antarmuka data latih menampilkan tabel data yang berisi nomor data, atribut jenis kelamin, usia, status, pekerjaan, pendapatan, tanggungan, OTR, DP, tenor, angsuran, serta kelas yang sebenarnya ketika pengguna menekan tombol "Data Latih" yang ditunjukkan pada Gambar 5.1.



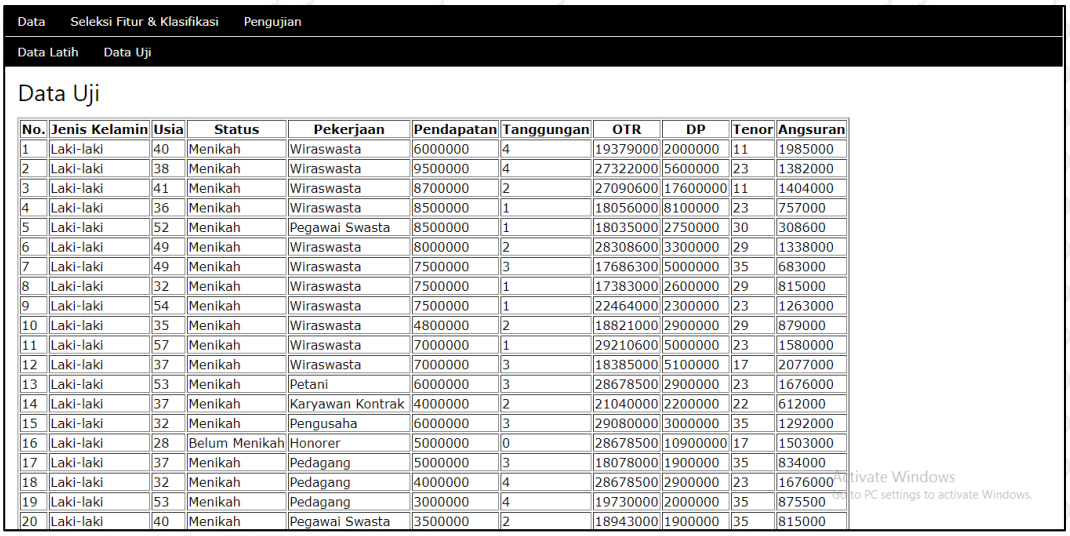
No.	Jenis Kelamin	Usia	Status	Pekerjaan	Pendapatan	Tanggungan	OTR	DP	Tenor	Angsuran	Klasifikasi
1	Perempuan	38	Menikah	Wiraswasta	30000000	4	26110600	2700000	35	1083000	Lancar
2	Perempuan	29	Menikah	Wiraswasta	20000000	2	17686300	2800000	17	1198000	Lancar
3	Perempuan	41	Menikah	Wiraswasta	7000000	3	35191400	10000000	29	1413000	Lancar
4	Laki-laki	30	Belum Menikah	Pegawai Swasta	7500000	0	27356700	5000000	26	1270300	Lancar
5	Perempuan	42	Menikah	Wiraswasta	6500000	1	22316000	2300000	29	1076000	Lancar
6	Laki-laki	52	Menikah	PNS	6000000	2	30731700	3100000	30	1273000	Lancar
7	Laki-laki	30	Belum Menikah	Pegawai Swasta	10000000	0	59951900	11000000	35	2203000	Lancar
8	Perempuan	28	Belum Menikah	Mahasiswi	6000000	0	17121000	5000000	29	717000	Lancar
9	Laki-laki	50	Widow	PNS	6000000	5	24860000	3700000	35	999000	Lancar
10	Laki-laki	35	Menikah	Wiraswasta	6000000	3	17121000	1800000	35	752000	Lancar
11	Laki-laki	41	Belum Menikah	Pegawai Swasta	6000000	0	28308600	17100000	11	1798000	Lancar
12	Laki-laki	67	Menikah	Wiraswasta	6000000	0	16051000	5000000	24	711300	Lancar
13	Laki-laki	51	Menikah	Pegawai Swasta	6000000	2	19285000	2900000	30	855200	Lancar
14	Laki-laki	45	Menikah	Pegawai Swasta	5800000	3	17635000	2700000	35	766400	Lancar
15	Laki-laki	54	Menikah	PNS	5500000	5	28078500	2900000	35	1162000	Lancar
16	Laki-laki	36	Menikah	Wiraswasta	5000000	2	18508200	1900000	27	896000	Lancar
17	Laki-laki	40	Menikah	Petani	5000000	2	16887000	3100000	29	773000	Lancar
18	Perempuan	27	Belum Menikah	Wiraswasta	5000000	0	22464000	9100000	29	845000	Lancar
19	Perempuan	51	Menikah	Wiraswasta	5000000	3	19535000	4000000	23	1013000	Lancar
20	Laki-laki	45	Menikah	Wiraswasta	5000000	2	18385000	8100000	23	760000	Lancar

Gambar 5.1. Implementasi Antarmuka Data Latih

5.4.1.2 Implementasi Antarmuka Data Uji

Antarmuka data uji menampilkan tabel data yang berisi nomor data, atribut jenis kelamin, usia, status, pekerjaan, pendapatan, tanggungan, OTR, DP,

tenor, angsuran, serta kelas yang sebenarnya ketika pengguna menekan tombol "Data Uji" yang ditunjukkan pada Gambar 5.2.



No.	Jenis Kelamin	Usia	Status	Pekerjaan	Pendapatan	Tanggungan	OTR	DP	Tenor	Angsuran
1	Laki-laki	40	Menikah	Wiraswasta	6000000	4	19379000	2000000	11	1985000
2	Laki-laki	38	Menikah	Wiraswasta	9500000	4	27322000	5600000	23	1382000
3	Laki-laki	41	Menikah	Wiraswasta	8700000	2	27090600	17600000	11	1404000
4	Laki-laki	36	Menikah	Wiraswasta	8500000	1	18056000	8100000	23	757000
5	Laki-laki	52	Menikah	Pegawai Swasta	8500000	1	18035000	2750000	30	308600
6	Laki-laki	49	Menikah	Wiraswasta	8000000	2	28308600	3300000	29	1338000
7	Laki-laki	49	Menikah	Wiraswasta	7500000	3	17686300	5000000	35	683000
8	Laki-laki	32	Menikah	Wiraswasta	7500000	1	17383000	2600000	29	815000
9	Laki-laki	54	Menikah	Wiraswasta	7500000	1	22464000	2300000	23	1263000
10	Laki-laki	35	Menikah	Wiraswasta	4800000	2	18821000	2900000	29	879000
11	Laki-laki	57	Menikah	Wiraswasta	7000000	1	29210600	5000000	23	1580000
12	Laki-laki	37	Menikah	Wiraswasta	7000000	3	18385000	5100000	17	2077000
13	Laki-laki	53	Menikah	Petani	6000000	3	28678500	2900000	23	1676000
14	Laki-laki	37	Menikah	Karyawan Kontrak	4000000	2	21040000	2200000	22	612000
15	Laki-laki	32	Menikah	Pengusaha	6000000	3	29080000	3000000	35	1292000
16	Laki-laki	28	Belum Menikah	Honorer	5000000	0	28678500	10900000	17	1503000
17	Laki-laki	37	Menikah	Pedagang	5000000	3	18078000	1900000	35	834000
18	Laki-laki	32	Menikah	Pedagang	4000000	4	28678500	2900000	23	1676000
19	Laki-laki	53	Menikah	Pedagang	3000000	4	19730000	2000000	35	875500
20	Laki-laki	40	Menikah	Pegawai Swasta	3500000	2	18943000	1900000	35	815000

Activate Windows
Go to PC settings to activate Windows.

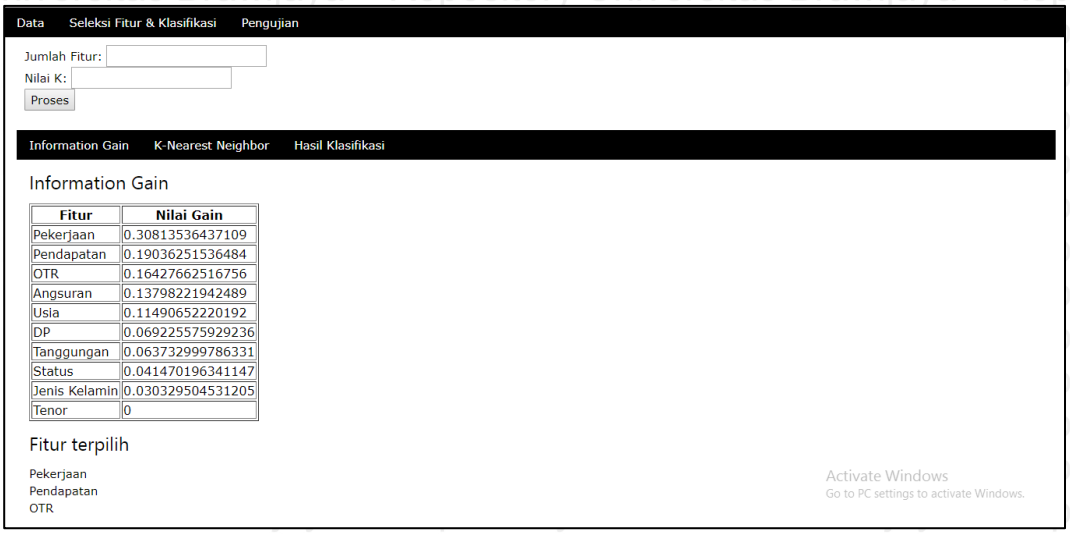
Gambar 5.2. Implementasi Antarmuka Data Uji

5.4.2 Implementasi Antarmuka Proses Seleksi Fitur dan Klasifikasi

Antarmuka kedua yaitu antarmuka dalam penerapan metode yang digunakan. Antarmuka ini akan menampilkan form masukan jumlah fitur dan nilai *K* serta terdapat pula tombol proses ketika pengguna menekan tombol "Seleksi Fitur & Klasifikasi". Dua inputan tersebut berguna untuk proses seleksi fitur dan klasifikasi. Hasil perhitungan metode *Information Gain* ditampilkan pada saat pengguna menekan tombol "Information Gain" dan hasil perhitungan metode *K-Nearest Neighbor* ditampilkan pada saat pengguna menekan tombol "K-Nearest Neighbor".

5.4.2.1 Implementasi Antarmuka *Information Gain*

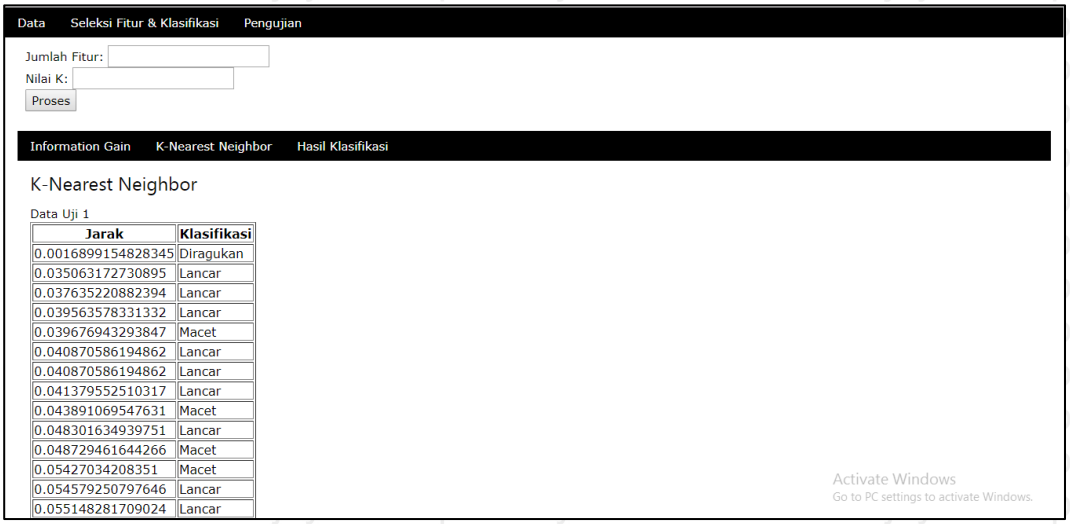
Antarmuka *Information Gain* menampilkan tabel perhitungan gain dari masing-masing atribut ketika pengguna menekan tombol proses. Tabel tersebut berisi nama-nama atribut beserta nilai *gain* dari atribut tersebut yang telah diurutkan. Gambar 5.3 menunjukkan penerapan antarmuka *Information Gain*.



Gambar 5.3. Implementasi Antarmuka *Information Gain*

5.4.2.2 Implementasi Antarmuka *K-Nearest Neighbor*

Antarmuka *K-Nearest Neighbor* menampilkan tabel hasil perhitungan jarak antara data latih dengan data uji ketika pengguna menekan tombol proses. Tabel tersebut berisi nilai jarak beserta label kelas dari data yang telah diurutkan dan pengimplementasiannya dapat dilihat pada Gambar 5.4.



Gambar 5.4. Implementasi Antarmuka *K-Nearest Neighbor*

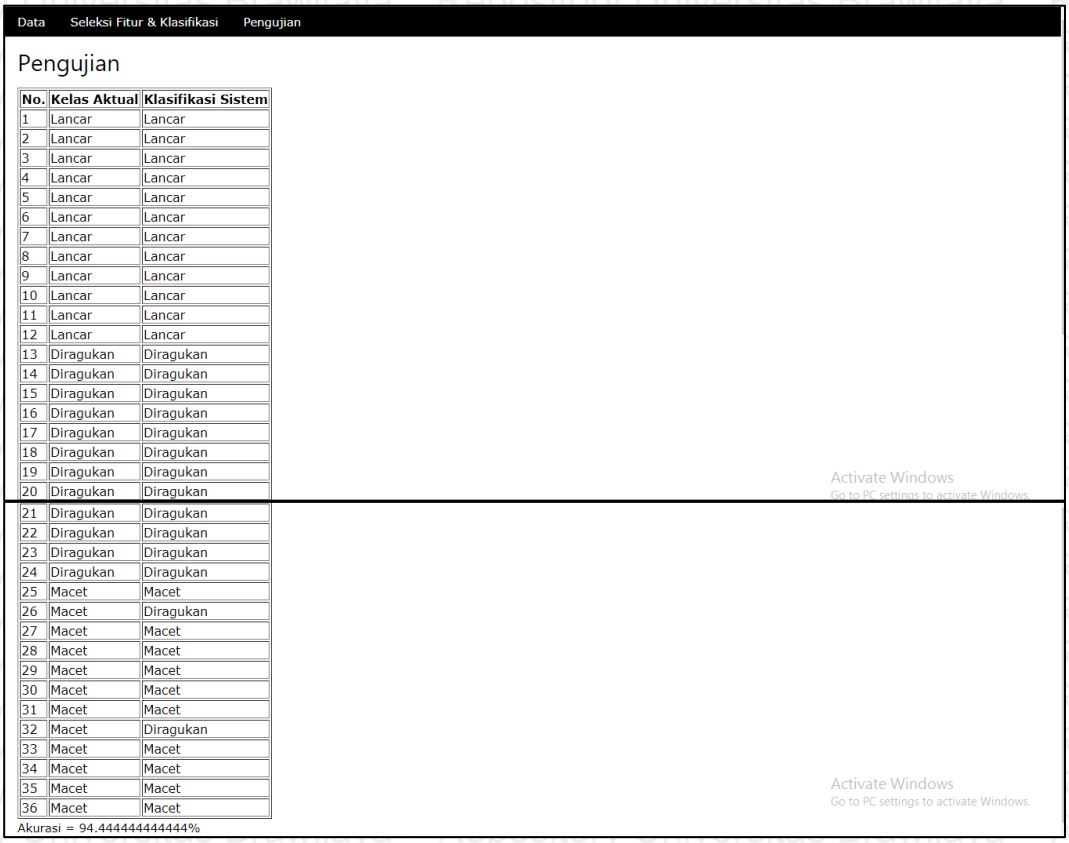
5.4.2.3 Implementasi Antarmuka Hasil Klasifikasi

Antarmuka hasil klasifikasi menunjukkan tabel yang berisi nomor data dan hasil klasifikasi yang merupakan prediksi dari sistem ketika pengguna menekan tombol “Hasil Klasifikasi” dan pengimplementasiannya dapat dilihat pada Gambar 5.5.

Gambar 5.5. Implementasi Antarmuka Hasil Klasifikasi

5.4.3 Implementasi Antarmuka Pengujian

Antarmuka pengujian menampilkan tabel yang berisi nomor data, kelas data yang sebenarnya dan hasil klasifikasi yang merupakan prediksi dari sistem ketika pengguna menekan tombol "Pengujian" dan pengimplementasiannya ditunjukkan pada Gambar 5.6.



Gambar 5.6. Implementasi Antarmuka Pengujian

BAB 6 PENGUJIAN DAN ANALISIS

6.1 Pengujian

Pada sub bab ini akan menjelaskan mengenai pengujian yang dilakukan pada penelitian. Penelitian yang dibuat akan diuji melalui dua skenario. Skenario yang pertama yaitu dengan menggunakan sebaran kelas seimbang pada data latih dan skenario kedua yaitu dengan menggunakan sebaran kelas tidak seimbang pada data latih. Pada skenario kedua dibagi lagi menjadi dua skenario untuk menguji tingkat akurasinya. Jumlah data yang dipakai sebanyak 180 data yang terdiri dari 144 data latih dan 36 data uji. Pembagian jumlah data tersebut diterapkan baik pada skenario pengujian pertama maupun skenario pengujian kedua. Pengujian dilakukan berdasarkan masukan pada jumlah fitur dimulai dari 2 sampai dengan 10 serta nilai *K* dimulai dari 3 sampai dengan 10.

6.1.1 Pengujian Sebaran Kelas Seimbang pada Data Latih

Data latih yang digunakan dalam proses pengujian untuk skenario ini sebanyak 144 data yang terdiri dari 48 data berlabel Lancar, 48 data berlabel Diragukan, dan 48 data berlabel Macet, sedangkan data uji sebanyak 36 data yang terdiri dari 12 data berlabel Lancar, 12 data berlabel Diragukan, dan 12 data berlabel Macet. Hasil pengujian terhadap sebaran kelas seimbang ditunjukkan pada Tabel 6.1.

Tabel 6.1. Hasil Pengujian Sebaran Kelas Seimbang

Jumlah Fitur	Akurasi(%)							
	K=3	*K*=4	*K*=5	*K*=6	*K*=7	*K*=8	*K*=9	*K*=10
2	83,33	80,56	83,33	80,56	77,78	83,33	86,11	83,33
3	77,78	94,44	88,89	88,89	83,33	83,33	80,56	83,33
4	83,33	80,56	77,78	80,56	86,11	80,56	80,56	80,56
5	69,44	72,22	75	69,44	72,22	63,89	66,67	61,11
6	66,67	66,67	63,89	66,67	66,67	69,44	55,56	61,11
7	58,33	58,33	55,56	61,11	66,67	61,11	66,67	61,11
8	61,11	63,89	55,56	61,11	66,67	66,67	69,44	66,67
9	50	50	50	52,78	61,11	52,78	52,78	52,78
10	41,67	41,67	41,67	55,56	44,44	55,56	50	50

Dilihat dari Tabel 6.1 akurasi terendah yang diperoleh sebesar 41,67% ketika menggunakan jumlah fitur 10 dengan nilai *K*=3, 4, dan 5. Sedangkan nilai akurasi tertinggi diperoleh ketika menggunakan jumlah fitur 3 dengan nilai *K*=4 sebesar 94,44%.

6.1.2 Pengujian Sebaran Kelas Tidak Seimbang pada Data Latih

Proses pengujian sebaran kelas tidak seimbang ini dilakukan dengan dua skenario. Perbedaan dari kedua skenario tersebut terletak pada pembagian jumlah kelas data latih yang digunakan. Pada skenario pertama data latih yang

Repository Universitas Brawijaya
REPOSITORY.UB.AC.ID
UNIVERSITAS BRAWIJAYA

digunakan terdiri dari kelas Lancar sebanyak 60 data, kelas Diragukan sebanyak 48 data, dan kelas Macet sebanyak 36 data. Sedangkan pada skenario kedua data latih yang digunakan terdiri dari kelas Lancar sebanyak 60 data, kelas Diragukan sebanyak 36 data, dan kelas Macet sebanyak 48 data. Jumlah data uji yang digunakan sama dengan jumlah data uji yang digunakan pada pengujian sebelumnya. Tabel 6.2 dan Tabel 6.3 menunjukkan hasil akurasi yang dihasilkan pada pengujian ini.

Tabel 6.2. Hasil Pengujian Sebaran Kelas Tidak Seimbang (Skenario 1)

Jumlah Fitur	Akurasi(%)							
	K=3	*K*=4	*K*=5	*K*=6	*K*=7	*K*=8	*K*=9	*K*=10
2	80,56	75	77,78	72,22	75	75	72,22	69,44
3	77,78	83,33	80,56	83,33	83,33	83,33	80,56	80,56
4	80,56	83,33	80,56	83,33	77,78	77,78	80,56	77,78
5	63,89	69,44	72,22	66,67	69,44	61,11	63,89	52,78
6	63,89	66,67	66,67	63,89	58,33	55,56	52,78	58,33
7	52,78	58,33	55,56	52,78	55,56	52,78	55,56	55,56
8	63,89	58,33	50	44,44	52,78	55,56	58,33	52,78
9	47,22	44,44	38,89	44,44	52,78	47,22	44,44	50
10	38,89	38,89	41,67	47,22	47,22	47,22	44,44	44,44

Dilihat dari Tabel 6.2 akurasi terendah yang diperoleh sebesar 38,89% ketika menggunakan jumlah fitur 9 dengan nilai *K*=5 dan jumlah fitur 10 dengan nilai *K*=3 dan 4. Sedangkan nilai akurasi tertinggi diperoleh ketika menggunakan jumlah fitur 3 dengan nilai *K*=4,6,7,8 dan jumlah fitur 4 dengan nilai *K*=4 dan 6 sebesar 83,33%.

Tabel 6.3. Hasil Pengujian Sebaran Kelas Tidak Seimbang (Skenario 2)

Jumlah Fitur	Akurasi(%)							
	K=3	*K*=4	*K*=5	*K*=6	*K*=7	*K*=8	*K*=9	*K*=10
2	77,78	69,44	75	75	75	75	69,44	72,22
3	77,78	83,33	83,33	83,33	83,33	80,56	77,78	83,33
4	77,78	80,56	77,78	83,33	83,33	83,33	83,33	83,33
5	69,44	75	72,22	75	72,22	66,67	63,89	61,11
6	61,11	69,44	61,11	66,67	61,11	58,33	50	55,56
7	58,33	55,56	52,78	58,33	61,11	63,89	55,56	52,78
8	55,56	55,56	50	52,78	55,56	55,56	52,78	50
9	44,44	47,22	44,44	47,22	50	52,78	50	47,22
10	38,89	36,11	33,33	44,44	50	50	47,22	44,44

Dilihat dari Tabel 6.3 akurasi terendah yang diperoleh sebesar 33,33% ketika jumlah fitur 10 dengan nilai *K*=5. Sedangkan nilai akurasi tertinggi diperoleh ketika jumlah fitur 3 dengan nilai *K*=4,5,6,7,10 dan jumlah fitur 4 dengan nilai *K*=6,7,8,9,10 sebesar 83,33%.

Repository Universitas Brawijaya
REPOSITORY.UB.AC.ID
UNIVERSITAS BRAWIJAYA

6.2 Analisis Hasil Pengujian

Berdasarkan skenario pengujian dengan menggunakan jumlah fitur mulai dari 2 sampai 10 dengan nilai *K* yang dimulai dari 3 sampai 10 terhadap jumlah data uji yang sama dan jumlah data latih yang sama dengan label kelas seimbang maupun tidak seimbang diperoleh hasil akurasi yang berbeda-beda. Nilai akurasi terendah dihasilkan pada pengujian dengan sebaran kelas tidak seimbang menggunakan jumlah fitur jumlah fitur 10 dengan nilai *K*=5 yaitu sebesar 33,33%, sedangkan pada pengujian dengan sebaran kelas seimbang menggunakan jumlah fitur 10 dengan nilai *K*=3, 4, dan 5 memperoleh nilai akurasi terendah sebesar 41,67%. Nilai akurasi tertinggi pada pengujian dengan sebaran kelas tidak seimbang yaitu sebesar 83,33% dengan menggunakan jumlah fitur 3 dengan nilai *K*=4,6,7,8 dan jumlah fitur 4 dengan nilai *K*=4 dan 6 pada skenario 1 serta menggunakan jumlah fitur 3 dengan nilai *K*=4,5,6,7,10 dan jumlah fitur 4 dengan nilai *K*=6,7,8,9,10 pada skenario 2. Pada pengujian dengan sebaran kelas seimbang menggunakan jumlah fitur 3 dengan nilai *K*=4 nilai akurasi tertinggi diperoleh sebesar 94,44%.

Dari hasil pengujian berdasarkan kedua skenario pada sub bab sebelumnya, fitur atau atribut yang dihasilkan memilki urutan yang sama. Urutan tersebut berdasarkan dari perangkingan terhadap besarnya nilai *gain* dari setiap atribut. Urutan atribut ini pula yang menjadi alasan dalam penentuan jumlah sebaran data yang digunakan dalam pengujian, baik pengujian dengan skenario 1 maupun skenario 2. Tabel 6.4 menunjukkan fitur dari *Information Gain*.

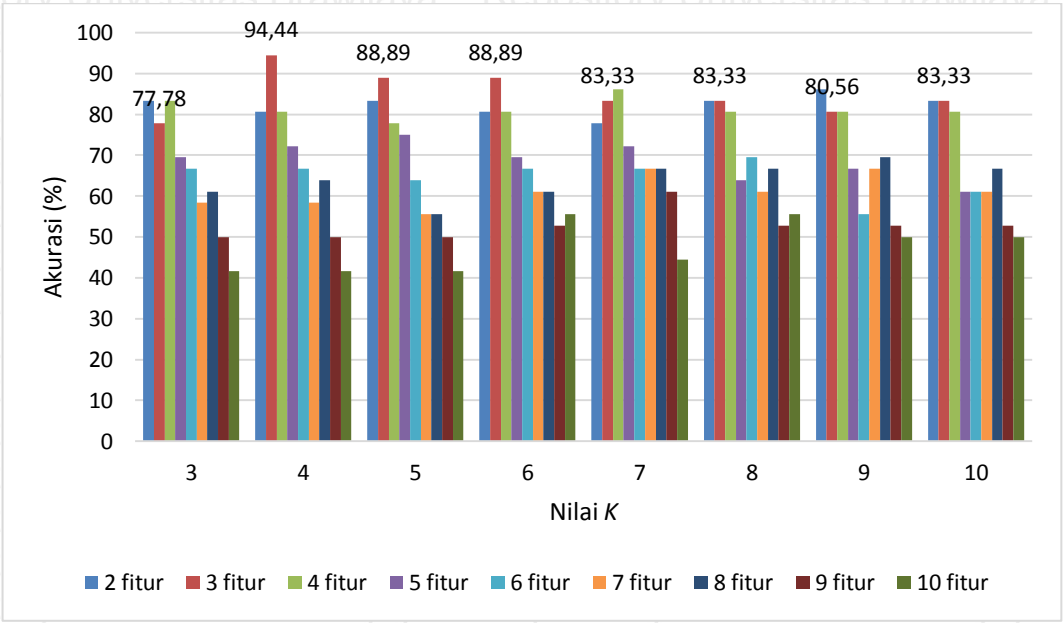
Tabel 6.4 Hasil Seleksi Fitur *Information Gain*

Jumlah Fitur	Fitur/atribut terpilih	
	Sebaran Kelas Seimbang	Sebaran Kelas Tidak Seimbang
2	Pekerjaan, pendapatan	Pekerjaan, pendapatan
3	Pekerjaan, pendapatan, OTR	Pekerjaan, pendapatan, OTR
4	Pekerjaan, pendapatan, OTR, angsuran	Pekerjaan, pendapatan, OTR, angsuran
5	Pekerjaan, pendapatan, OTR, angsuran, usia	Pekerjaan, pendapatan, OTR, angsuran, usia
6	Pekerjaan, pendapatan, OTR, angsuran, usia, DP	Pekerjaan, pendapatan, OTR, angsuran, usia, DP
7	Pekerjaan, pendapatan, OTR, angsuran, usia, DP, tanggungan	Pekerjaan, pendapatan, OTR, angsuran, usia, DP, tanggungan
8	Pekerjaan, pendapatan, OTR, angsuran, usia, DP, tanggungan, status	Pekerjaan, pendapatan, OTR, angsuran, usia, DP, tanggungan, status

Repository Universitas Brawijaya
REPOSITORY.UB.AC.ID
UNIVERSITAS BRAWIJAYA

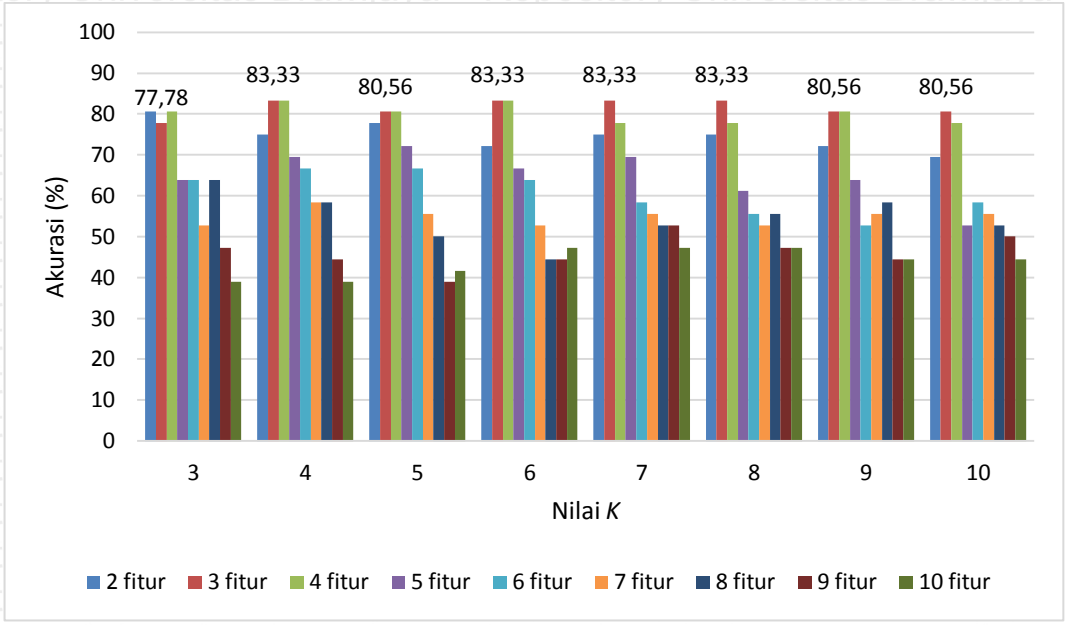
Jumlah Fitur	Fitur/atribut terpilih	
	Sebaran Kelas Seimbang	Sebaran Kelas Tidak Seimbang
9	Pekerjaan, pendapatan, OTR, angsuran, usia, DP, tanggungan, status, jenis kelamin	Pekerjaan, pendapatan, OTR, angsuran, usia, DP, tanggungan, status, jenis kelamin
10	Pekerjaan, pendapatan, OTR, angsuran, usia, DP, tanggungan, status, jenis kelamin, tenor	Pekerjaan, pendapatan, OTR, angsuran, usia, DP, tanggungan, status, jenis kelamin, tenor

Dalam penelitian ini, baik jumlah fitur maupun nilai *K* dapat memberikan pengaruh pada proses klasifikasi. Jumlah fitur akan menentukan fitur atau atribut apa saja yang akan digunakan untuk menghitung jarak kedekatan antar data pada algoritme *K-Nearest Neighbor*. Nilai *K* dapat memberikan pengaruh dalam penentuan kelas hasil prediksi dari sistem. Jumlah kelas data pada penelitian ini berjumlah 3 atau ganjil sehingga nilai *K* dengan kategori genap lebih berpengaruh dalam penentuan kelas agar mengurangi kemungkinan munculnya jumlah kelas kelas yang sama dalam hasil klasifikasi. Gambar 6.1 menunjukkan grafik tingkat akurasi dari pengujian sebaran kelas seimbang. Gambar 6.2 dan Gambar 6.3 menunjukkan grafik tingkat akurasi dari pengujian sebaran kelas tidak seimbang.



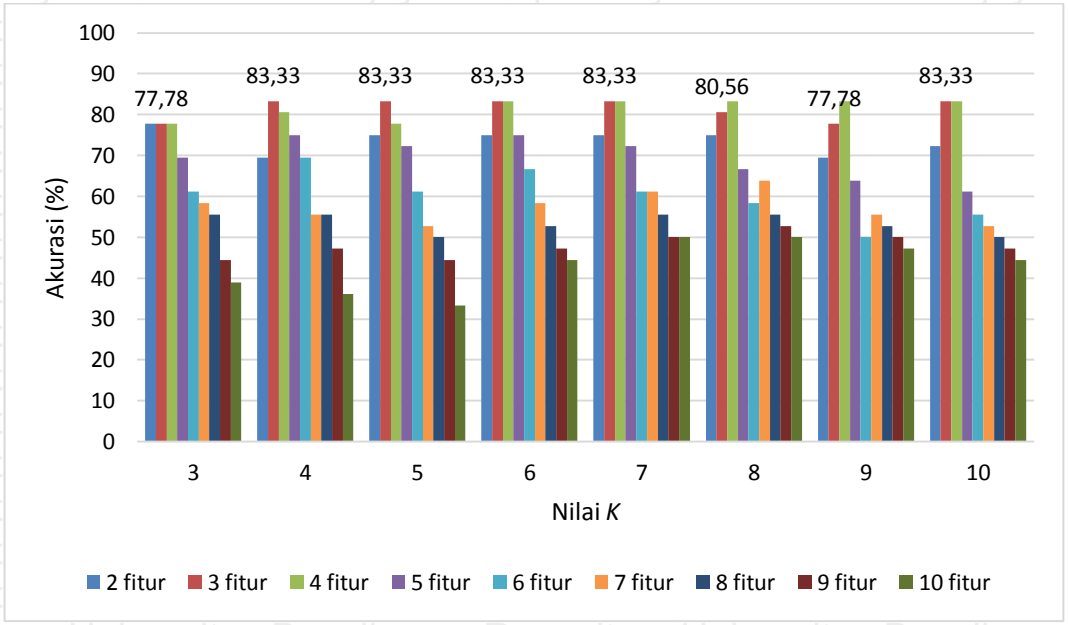
Gambar 6.1 Grafik Hasil Pengujian Sebaran Kelas Seimbang

Berdasarkan Gambar 6.1 akurasi tertinggi diperoleh pada saat jumlah fitur atau atribut 3 dengan nilai *K*=4 sebesar 94,44%. Tiga fitur tersebut adalah pekerjaan, pendapatan, dan OTR.



Gambar 6.2 Grafik Hasil Pengujian Sebaran Kelas Tidak Seimbang Skenario 1

Dilihat dari Gambar 6.2 akurasi tertinggi diperoleh pada saat jumlah fitur atau atribut 3 dan nilai *K*=4, 6,7,8 dan jumlah fitur 4 dengan nilai *K*=4 dan 6 sebesar 83,33%. Empat fitur tersebut pekerjaan, pendapatan, OTR, dan angsuran.



Gambar 6.3 Grafik Hasil Pengujian Sebaran Kelas Tidak Seimbang Skenario 2

Dilihat dari Gambar 6.3 nilai akurasi tertinggi diperoleh pada saat jumlah fitur 3 dengan nilai *K*=4,5,6,7,10 dan jumlah fitur 4 dengan nilai *K*=6,7,8,9,10 sebesar 83,33%. Empat fitur tersebut pekerjaan, pendapatan, OTR, dan angsuran.

Dilihat dari Gambar 6.1 sampai 6.3, akurasi terendah dihasilkan ketika pengujian dengan menggunakan 10 fitur atau semua fitur yaitu fitur pekerjaan,

pendapatan, OTR, angsuran, usia, DP, tanggungan, status, jenis kelamin, dan tenor, baik pada pengujian dengan skenario 1 maupun skenario 2. Hal tersebut berkaitan dengan salah satu kekurangan dari algoritme *K-Nearest Neighbor* yang rentan terhadap dimensionalitas yang tinggi. Semakin banyak fitur, ruang yang dapat ditempati suatu data semakin besar sehingga semakin besar pula kemungkinan suatu data masuk ke dalam klasifikasi label kelas tertentu yang sebenarnya jauh dari label kelas aktualnya. Dengan menggunakan 10 fitur berarti perhitungan jarak dilakukan dengan 10 dimensi yang berarti jarak semakin besar sehingga kedekatan antara dua instance semakin jauh. Berdasarkan Gambar 6.1 sampai 6.3, terlihat bahwa akurasi yang tertinggi terdapat ketika menggunakan 3 fitur dan akurasi terendah pada saat menggunakan 10 fitur. Perbandingan akurasi tersebut dapat dilihat pada Tabel 6.5.

Tabel 6.5. Perbandingan Akurasi Kelas Seimbang dan Tidak Seimbang dengan 10 fitur dan 3 fitur

Nilai *K*	Sebaran Kelas Tidak Seimbang		Sebaran Kelas Seimbang	
	Akurasi dengan 10 fitur (Tanpa *Information Gain*)	Akurasi dengan 3 fitur (Dengan *Information Gain*)	Akurasi dengan 10 fitur (Tanpa *Information Gain*)	Akurasi dengan 3 fitur (Dengan *Information Gain*)
3	38,89	77,78	41,67	77,78
4	38,89	83,33	41,67	94,44
5	41,67	80,56	41,67	88,89
6	47,22	83,33	55,56	88,89
7	47,22	83,33	44,44	83,33
8	47,22	83,33	55,56	83,33
9	44,44	80,56	50	80,56
10	44,44	80,56	50	83,33

Tabel 6.5 memperlihatkan bahwa penggunaan seleksi fitur *Information Gain* menghasilkan nilai akurasi yang lebih baik dibandingkan tanpa menggunakan seleksi fitur *Information Gain*. Tanpa menggunakan seleksi fitur *Information Gain* akurasi yang dihasilkan oleh sistem pada saat nilai *K*=4 menunjukkan hasil yang kurang baik yaitu sebesar 38,89% pada sebaran kelas tidak seimbang dan sebesar 41,67% pada sebaran kelas seimbang. Berbeda dengan menggunakan seleksi fitur *Information Gain* dihasilkan akurasi yang cukup baik yaitu sebesar 83,33% pada sebaran kelas tidak seimbang dan sebesar 94,44% pada sebaran kelas seimbang dengan menggunakan 3 fitur. Tiga fitur terebut adalah pekerjaan, pendapatan, dan OTR yang dapat dilihat pada Tabel 6.4. Fitur-fitur terpilih tersebut memiliki urutan nilai gain terbesar atau memiliki nilai gain lebih dari 0,1. Nilai *information gain* untuk setiap fitur sangat dipengaruhi oleh nilai data yang ada didalamnya. Semakin banyak variasi nilai yang ada maka semakin tinggi kemungkinan fitur memiliki nilai *information gain* yang tinggi. Hal tersebut terbukti pada fitur pekerjaan yang memiliki variasi 19 kategori sehingga menghasilkan nilai *gain* tertinggi dalam pengujian. Kesimpulan yang didapat setelah dilakukannya

pengujian bahwa pengurangan fitur sangat berpengaruh pada hasil klasifikasi dikarenakan seleksi fitur dapat mengurangi fitur yang tidak relevan terhadap kelas target atau yang tidak terlalu memberikan pengaruh dan mengurangi dimensi pada data sehingga dapat meningkatkan akurasi pada algoritme klasifikasi.

BAB 7 PENUTUP

7.1 Kesimpulan

Berdasarkan hasil yang diperoleh dari pengujian yang telah dilakukan pada penelitian ini, beberapa kesimpulan yang diperoleh adalah:

1. Pengujian pada penelitian ini dilakukan dengan sebaran kelas seimbang dan dengan sebaran kelas tidak seimbang. Nilai akurasi tertinggi diperoleh pada pengujian dengan sebaran kelas seimbang sebesar 94,44%, sedangkan pada pengujian dengan sebaran kelas tidak seimbang dihasilkan nilai akurasi tertinggi sebesar 83,33%.
2. Seleksi fitur *Information Gain* memberikan pengaruh terhadap klasifikasi tingkat kelancaran pembayaran kredit kendaraan yaitu peningkatan akurasi. Akurasi dengan seleksi fitur menghasilkan akurasi lebih baik jika dibandingkan dengan menggunakan 10 fitur atau tanpa seleksi fitur dan fitur yang menghasilkan akurasi tertinggi sejumlah tiga fitur yaitu pekerjaan, pendapatan dan OTR. Tiga fitur tersebut merupakan fitur yang menghasilkan nilai *gain* lebih dari 0,1. Besarnya nilai *gain* dalam suatu atribut menentukan besarnya pengaruh atribut tersebut dalam proses klasifikasi.

7.2 Saran

Berdasarkan kesimpulan yang dijabarkan pada sub bab sebelumnya, terdapat beberapa saran untuk pengembangan penelitian selanjutnya yaitu:

1. Melakukan perbandingan dengan menerapkan metode seleksi fitur lain sehingga dapat diketahui apakah metode seleksi fitur lain tersebut lebih baik dari pada *Information Gain*. Salah satunya dengan menggunakan analisis korelasi.
2. Menerapkan metode *Naïve Bayes* karena data yang digunakan dalam penelitian merupakan data nominal atau data kategori.
3. Melakukan pembobotan pada kelas tetangga sehingga kelas dengan jumlah data latih yang sedikit dapat memiliki kesempatan untuk menjadi hasil klasifikasi.
4. Penambahkan data latih untuk keberagaraman data di setiap kelas sehingga data dapat diklasifikasi lebih baik.

Repository Universitas Brawijaya
REPOSITORY.UB.AC.ID
UNIVERSITAS BRAWIJAYA

DAFTAR REFERENSI

Aini, S. H. A., Sari, Y. A. & Arwan, A., 2018. Seleksi Fitur Information Gain untuk Klasifikasi Penyakit Jantung Menggunakan Kombinasi Metode K-Nearest Neighbor dan Naïve Bayes. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer,* Volume 2, pp. 2546-2554.

Arifin, M., 2015. IG-KNN untuk Prediksi Customer Churn Telekomunikasi. *Jurnal Teknik Industri, Mesin, Elektro, dan Ilmu Komputer,* Volume 6, pp. 1-10.

Astiko, 1996. *Manajemen Perkreditan.* Yogyakarta: Andi Offset.

Azhagusundari, B. & Thanamani, A. S., 2013. Feature Selection based on Information Gain. *International Journal of Innovative Technology and Exploring Engineering,* 2(2).

Bahri, R. S. & Maliki, I., 2012. Perbandingan Algoritma Template Matching dan Feature Extraction Pada Optical Character Recognition. *Jurnal Komputer dan Informatika (KOMPUTA),* Volume 1.

Bramer, M., 2007. *Principles of Data Mining.* London: Springer.

Chormunge, S. & Jena, S., 2016. Efficient Feature Subset Selection Algorithm for High Dimensional Data. *International Journal of Electrical and Computer Engineering,* Volume 6, pp. 1880-1888.

Gorunescu, F., 2011. *Data Mining: Concepts, models, and techniques.* Springer-Verlag Berlin Heidelberg: Intelligent Systems References Library.

Heryono & Kardianawati, A., 2018. Impelmentasi Metode Naive Bayes untuk Klasifikasi Kredit Motor. *Journal Of Information System,* Volume 3, pp. 10-21.

Hsu, H.-H., 2006. *Advanced Data mining Technologies in Bioinformatics.* United States of America: Idea Group Publishing.

Ivandari, Chasanah, T. T., Binabar, S. W. & Karomi, M. A. A., 2017. Data Attribute Selection with Information Gain to Improve Credit Approval Classification Performance using K-Nearest Neighbor Algorithm. *International Journal of Islamic Business and Economics,* Volume 1, pp. 13-22.

Junaedi, H., Budianto, H., Maryati, I. & Melani, Y., 2011. *Data Tranformation pada Data Mining.* Surabaya, IdeaTech.

Kamber, M. & Han, J., 2006. *Data Mining Concepts And Techniques 2nd Edition.* San Fransisco: Elsevier.

Larose, D. T., 2005. *Discovering Knowledge in Data : An Introduction to Data Mining.* New Jersey: John Willey & Sons.

Leidiana, H., 2013. Penerapan Algoritma K-Nearest Neighbor untuk Penentuan Resiko Kredit Kepemilikan Kendaraan Bermotor. *Jurnal Penelitian Ilmu Komputer, System Embedded & Logic,* Volume 1, pp. 65-76.

Poonguzhali, E., Vijayalakshmi, M., Shamily, K. & Priyadarshini, V., 2014. Hybrid Feature Selection Algorithm for High Dimensional Database. *International Journal Of Engineering Trends And Technology (IJETT),* Volume 8, pp. 490-494.

Rivai, V., 2006. *Credit Management Handbook.* Jakarta: Raja Grafindo Persada.

Santoso, B., 2007. *Data Mining Teknik Pemanfaatan Data untuk Keperluan Bisnis.* Yogyakarta: Graha Ilmu.

Shaltout, N. A., El-Hefnawi, M., Rafea, A. & Moustafa, A., 2014. *Information Gain as a Feature Selection Method for the Efficient Classification of Influenza Based on Viral Hosts.* London,U.K, WCE.

Witten, I. H., Frank, E. & Hall, M. A., 2011. *Data Mining: Practical Machine Learning Tools and Techniques.* Burlington: Morgan Kaufmann.

Repository Universitas Brawijaya
REPOSITORY.UB.AC.ID
UNIVERSITAS BRAWIJAYA

LAMPIRAN A DATA KREDIT KENDARAAN

No	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	Klasifikasi
1	Perempuan	38	Menikah	Wiraswasta	30000000	4	26110600	2700000	35	1083000	Lancar
2	Perempuan	29	Menikah	Wiraswasta	20000000	2	17686300	2800000	17	1198000	Lancar
3	Laki-laki	30	Belum Menikah	Pegawai Swasta	10000000	0	59951900	11000000	35	2203000	Lancar
4	Laki-laki	38	Menikah	Wiraswasta	9500000	4	27322000	5600000	23	1382000	Lancar
5	Laki-laki	41	Menikah	Wiraswasta	8700000	2	27090600	17600000	11	1404000	Lancar
6	Laki-laki	36	Menikah	Wiraswasta	8500000	1	18056000	8100000	23	757000	Lancar
7	Laki-laki	52	Menikah	Pegawai Swasta	8500000	1	18035000	2750000	30	308600	Lancar
8	Laki-laki	49	Menikah	Wiraswasta	8000000	2	28308600	3300000	29	1338000	Lancar
9	Laki-laki	49	Menikah	Wiraswasta	7500000	3	17686300	5000000	35	683000	Lancar
10	Laki-laki	32	Menikah	Wiraswasta	7500000	1	17383000	2600000	29	815000	Lancar
11	Laki-laki	54	Menikah	Wiraswasta	7500000	1	22464000	2300000	23	1263000	Lancar
12	Laki-laki	30	Belum Menikah	Pegawai Swasta	7500000	0	27356700	5000000	26	1270300	Lancar
13	Laki-laki	57	Menikah	Wiraswasta	7000000	1	29210600	5000000	23	1580000	Lancar
14	Laki-laki	37	Menikah	Wiraswasta	7000000	3	18385000	5100000	17	2077000	Lancar
15	Perempuan	41	Menikah	Wiraswasta	7000000	3	35191400	10000000	29	1413000	Lancar
16	Laki-laki	49	Menikah	Wiraswasta	7000000	3	16647000	2500000	23	933600	Lancar
17	Perempuan	42	Menikah	Wiraswasta	6500000	1	22316000	2300000	29	1076000	Lancar
18	Laki-laki	52	Menikah	PNS	6000000	2	30731700	3100000	30	1273000	Lancar
19	Perempuan	28	Belum Menikah	Mahasiswi	6000000	0	17121000	5000000	29	717000	Lancar
20	Laki-laki	40	Menikah	Wiraswasta	6000000	4	19379000	2000000	11	1985000	Lancar
21	Laki-laki	50	Widow	PNS	6000000	5	24860000	3700000	35	999000	Lancar
22	Laki-laki	35	Menikah	Wiraswasta	6000000	3	17121000	1800000	35	752000	Lancar

Repository Universitas Brawijaya Repository
UNIVERSITAS BRAWIJAYA
REPOSITORY.UB.AC.ID

No		A2	A3	A4	A5	A6	A7	A8	A9	A10	Klasifikasi
23	Laki-laki	41	Belum Menikah	Pegawai Swasta	6000000	0	28308600	17100000	11	1798000	Lancar
24	Laki-laki	67	Menikah	Wiraswasta	6000000	0	16051000	5000000	24	711300	Lancar
25	Laki-laki	51	Menikah	Pegawai Swasta	6000000	2	19285000	2900000	30	855200	Lancar
26	Laki-laki	45	Menikah	Pegawai Swasta	5800000	3	17635000	2700000	35	766400	Lancar
27	Laki-laki	54	Menikah	PNS	5500000	5	28078500	2900000	35	1162000	Lancar
28	Laki-laki	36	Menikah	Wiraswasta	5000000	2	18508200	1900000	27	896000	Lancar
29	Laki-laki	40	Menikah	Petani	5000000	2	16887000	3100000	29	773000	Lancar
30	Perempuan	27	Belum Menikah	Wiraswasta	5000000	0	22464000	9100000	29	845000	Lancar
31	Perempuan	51	Menikah	Wiraswasta	5000000	3	19535000	4000000	23	1013000	Lancar
32	Laki-laki	45	Menikah	Wiraswasta	5000000	2	18385000	8100000	23	760000	Lancar
33	Laki-laki	38	Menikah	Wiraswasta	5000000	2	27090600	9000000	23	1271000	Lancar
34	Laki-laki	52	Menikah	Wiraswasta	5000000	1	18385000	4900000	22	905000	Lancar
35	Laki-laki	35	Menikah	Wiraswasta	5000000	4	17383000	5000000	35	669000	Lancar
36	Laki-laki	50	Menikah	Wiraswasta	5000000	2	16678000	5100000	29	693000	Lancar
37	Laki-laki	54	Menikah	Wiraswasta	5000000	2	18340000	2500000	29	866000	Lancar
38	Laki-laki	45	Menikah	PNS	5000000	2	22016000	2300000	29	1060000	Lancar
39	Laki-laki	37	Menikah	Karyawan swasta	5000000	2	22316000	2300000	23	1254000	Lancar
40	Laki-laki	41	Menikah	Wiraswasta	5000000	4	17121000	2000000	35	746000	Lancar
41	Laki-laki	48	Menikah	Wiraswasta	5000000	1	18721000	1900000	29	909000	Lancar
42	Laki-laki	43	Menikah	Wiraswasta	5000000	1	17342000	4000000	26	764600	Lancar
43	Perempuan	46	Menikah	pegawai BUMN	5000000	3	17056700	2600000	23	944000	Lancar
44	Laki-laki	32	Menikah	Pegawai Swasta	5000000	2	24490600	2500000	29	1179000	Lancar
45	Laki-laki	66	Menikah	Wiraswasta	5000000	1	26426700	5300000	32	1042700	Lancar
46	Laki-laki	35	Menikah	Wiraswasta	4800000	2	18821000	2900000	29	879000	Lancar

UNIVERSITAS BRAWIJAYA
REPOSITORY.UB.AC.ID

No	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	Klasifikasi
47	Laki-laki	33	Menikah	Pegawai Swasta	4800000	1	16451000	2500000	26	804500	Lancar
48	Perempuan	61	Menikah	Wiraswasta	4500000	1	17101000	1800000	11	1804000	Lancar
49	Laki-laki	35	Menikah	Wiraswasta	4500000	4	18658200	1900000	30	820000	Lancar
50	Laki-laki	38	Menikah	PNS	4500000	3	17371000	2200000	30	750000	Lancar
51	Perempuan	51	Widow	PNS	4500000	2	24860000	4300000	29	1136000	Lancar
52	Laki-laki	37	Menikah	Karyawan Swasta	4500000	2	17592000	4000000	35	708000	Lancar
53	Laki-laki	39	Menikah	Wiraswasta	4500000	0	17686300	5000000	17	1066000	Lancar
54	Laki-laki	37	Menikah	Wiraswasta	4500000	3	18056000	3700000	35	738000	Lancar
55	Laki-laki	49	Menikah	Wiraswasta	4500000	1	27460500	5000000	35	1073000	Lancar
56	Laki-laki	45	Menikah	Wiraswasta	4500000	2	16678000	3500000	23	882000	Lancar
57	Laki-laki	42	Menikah	Wiraswasta	4500000	3	16287000	3300000	29	735600	Lancar
58	Perempuan	39	Menikah	Wiraswasta	4500000	3	16983000	3400000	20	941900	Lancar
59	Perempuan	31	Belum Menikah	Karyawan swasta	4500000	0	18074200	2100000	29	867000	Lancar
60	Laki-laki	38	Menikah	Wiraswasta	4500000	3	17383000	1800000	35	765000	Lancar
61	Laki-laki	38	Menikah	Wiraswasta	4500000	1	18135000	5000000	24	816200	Lancar
62	Laki-laki	25	Belum Menikah	Wiraswasta	4500000	0	18035000	2750000	30	808600	Lancar
63	Laki-laki	44	Menikah	Pegawai Swasta	4500000	1	16806700	3400000	30	698300	Lancar
64	Perempuan	37	Menikah	Karyawan Swasta	4250000	1	27460500	4800000	29	1249000	Lancar
65	Laki-laki	21	Belum Menikah	Wiraswasta	4000000	0	28308600	3300000	29	1338000	Lancar
66	Perempuan	51	Menikah	PNS	4000000	3	18943000	1900000	27	930000	Lancar
67	Laki-laki	30	Belum Menikah	Karyawan Swasta	4000000	0	27090600	5000000	23	1451000	Lancar
68	Laki-laki	40	Menikah	Wiraswasta	4000000	4	16701000	3100000	29	763000	Lancar
69	Perempuan	32	Belum Menikah	Wiraswasta	4000000	0	27090600	7200000	29	1145000	Lancar
70	Laki-laki	45	Menikah	Karyawan Swasta	4000000	1	30787600	10100000	29	1232000	Lancar

UNIVERSITAS BRAWIJAYA
REPOSITORY.UB.AC.ID

No		A2	A3	A4	A5	A6	A7	A8	A9	A10	Klasifikasi
71	Laki-laki	37	Menikah	PNS	4000000	3	27460500	5600000	23	1447000	Lancar
72	Perempuan	40	Menikah	Wiraswasta	4000000	2	16678000	6900000	29	630000	Lancar
73	Laki-laki	30	Menikah	Wiraswasta	4000000	1	19091000	2600000	32	790000	Lancar
74	Perempuan	41	Menikah	Wiraswasta	4000000	2	18427800	6000000	35	627600	Lancar
75	Laki-laki	35	Menikah	Wiraswasta	4000000	1	18074000	1900000	35	793000	Lancar
76	Laki-laki	34	Menikah	PNS	4000000	1	18721000	1900000	35	795000	Lancar
77	Laki-laki	37	Menikah	Wiraswasta	4000000	2	17436300	3500000	27	776600	Lancar
78	Laki-laki	37	Menikah	Wiraswasta	4000000	1	18135000	4000000	26	815900	Lancar
79	Perempuan	48	Menikah	PNS	4000000	4	15901000	2500000	35	676600	Lancar
80	Perempuan	29	Menikah	Pegawai Swasta	4000000	0	16051000	3250000	27	715900	Lancar
81	Laki-laki	56	Menikah	PNS	4000000	1	17342000	2650000	30	778400	Lancar
82	Perempuan	28	Belum Menikah	Karyawan Swasta	3800000	0	18056000	3100000	29	831000	Lancar
83	Laki-laki	40	Menikah	Wiraswasta	3800000	3	21566000	5000000	35	798600	Lancar
84	Perempuan	39	Menikah	Wiraswasta	3750000	3	18821000	8900000	35	590000	Lancar
85	Laki-laki	23	Belum Menikah	Wiraswasta	3600000	0	16451000	3500000	30	675800	Lancar
86	Laki-laki	24	Belum Menikah	Wiraswasta	3500000	0	19043000	2000000	30	815000	Lancar
87	Perempuan	30	Menikah	PNS	3500000	1	17686300	4000000	23	920000	Lancar
88	Perempuan	36	Menikah	Wiraswasta	3500000	0	16806700	5000000	26	681600	Lancar
89	Laki-laki	50	Widow	Wiraswasta	3500000	0	21596000	3400000	23	1162000	Lancar
90	Perempuan	32	Belum Menikah	Wiraswasta	3500000	0	19780000	2000000	35	881000	Lancar
91	Laki-laki	27	Belum Menikah	Karyawan Swasta	3500000	0	17121000	3800000	23	862000	Lancar
92	Laki-laki	29	Belum Menikah	Wiraswasta	3500000	0	21266000	4300000	35	816400	Lancar
93	Laki-laki	25	Belum Menikah	Pegawai Swasta	3500000	0	18176000	1900000	11	1899000	Lancar
94	Laki-laki	31	Belum Menikah	PNS	3200000	0	18821000	3900000	35	740000	Lancar

Repository Universitas Brawijaya
REPOSITORY.UB.AC.ID
UNIVERSITAS BRAWIJAYA

No		A2	A3	A4	A5	A6	A7	A8	A9	A10	Klasifikasi
95	Perempuan	46	Menikah	Wiraswasta	3000000	2	16051000	2450000	32	689200	Lancar
96	Perempuan	53	Menikah	PNS	3000000	0	17201000	1800000	11	1816000	Lancar
97	Laki-laki	33	Menikah	Wiraswasta	3000000	1	17436300	5000000	32	635200	Lancar
98	Perempuan	29	Menikah	Pegawai Swasta	3000000	1	15901000	1700000	21	982000	Lancar
99	Perempuan	25	Belum Menikah	Wiraswasta	3000000	0	16806700	3000000	30	733800	Lancar
100	Laki-laki	53	Menikah	Pegawai Swasta	3000000	1	19535000	4000000	23	1013000	Lancar
101	Laki-laki	61	Menikah	Supir	2800000	2	16806700	3400000	30	698300	Lancar
102	Perempuan	43	Menikah	Pegawai Swasta	2700000	1	18035000	2750000	30	808600	Lancar
103	Perempuan	26	Belum Menikah	Honorer	2500000	0	20883100	2100000	30	884000	Lancar
104	Laki-laki	24	Belum Menikah	Pegawai Swasta	2500000	0	17635000	3550000	35	725500	Lancar
105	Laki-laki	62	Menikah	Petani	11000000	2	28678500	17700000	11	1345000	Diragukan
106	Laki-laki	40	Menikah	Wiraswasta	10500000	2	16901000	1700000	30	747000	Diragukan
107	Laki-laki	47	Menikah	PNS	10000000	3	25818000	2600000	27	1323000	Diragukan
108	Perempuan	49	Belum Menikah	Wiraswasta	10000000	0	19400000	2000000	34	873000	Diragukan
109	Laki-laki	47	Menikah	Wiraswasta	10000000	1	25248000	2600000	35	1127000	Diragukan
110	Laki-laki	35	Menikah	PNS	9000000	2	28678500	2900000	35	1279000	Diragukan
111	Perempuan	43	Widow	PNS	8000000	4	29086000	3000000	23	1696000	Diragukan
112	Laki-laki	28	Menikah	Wiraswasta	8000000	2	26840000	4750000	20	1508500	Diragukan
113	Laki-laki	32	Menikah	Wiraswasta	8000000	4	16051000	3250000	24	787700	Diragukan
114	Laki-laki	28	Belum Menikah	Wiraswasta	7500000	0	19400000	2000000	30	863000	Diragukan
115	Laki-laki	37	Belum Menikah	Kontraktor	7000000	0	17235000	1800000	11	1844000	Diragukan
116	Perempuan	56	Widow	Pedagang	7000000	0	19836000	2000000	35	884000	Diragukan
117	Laki-laki	47	Menikah	Pedagang	7000000	9	17722000	1800000	34	811000	Diragukan
118	Laki-laki	37	Menikah	PNS	7000000	3	16450000	4000000	14	1188200	Diragukan

Repository Universitas Brawijaya
REPOSITORY.UB.AC.ID
UNIVERSITAS BRAWIJAYA

No		A2	A3	A4	A5	A6	A7	A8	A9	A10	Klasifikasi
119	Laki-laki	40	Menikah	Wiraswasta	6500000	3	16701000	6300000	35	599000	Diragukan
120	Perempuan	54	Widow	Karyawan Swasta	6100000	0	15878600	2400000	26	801200	Diragukan
121	Laki-laki	53	Menikah	Petani	6000000	3	28678500	2900000	23	1676000	Diragukan
122	Perempuan	57	Belum Menikah	Petani	6000000	0	39357000	4000000	22	2266000	Diragukan
123	Laki-laki	25	Menikah	Pedagang	6000000	1	28678500	10100000	23	1208200	Diragukan
124	Perempuan	26	Belum Menikah	Pegawai Swasta	6000000	0	28308600	2900000	29	1358000	Diragukan
125	Perempuan	52	Widow	Pedagang	6000000	0	17194000	1800000	34	785000	Diragukan
126	Laki-laki	41	Menikah	Karyawan Kontrak	6000000	3	17188000	1800000	32	784000	Diragukan
127	Laki-laki	32	Menikah	Pengusaha	6000000	3	29080000	3000000	35	1292000	Diragukan
128	Laki-laki	49	Menikah	Buruh	6000000	9	25360600	2600000	35	1132000	Diragukan
129	Laki-laki	32	Menikah	Polisi	6000000	1	27460500	5600000	29	1221000	Diragukan
130	Perempuan	32	Menikah	Ibu Rumah Tangga	6000000	3	19836000	2000000	35	884000	Diragukan
131	Laki-laki	56	Menikah	PNS	6000000	1	30731700	5100000	29	1374000	Diragukan
132	Laki-laki	45	Menikah	PNS	6000000	4	27708600	2800000	35	1150000	Diragukan
133	Perempuan	28	Menikah	Wiraswasta	6000000	2	19300000	2000000	29	984000	Diragukan
134	Perempuan	36	Widow	Wiraswasta	5900000	0	16551000	1700000	30	750000	Diragukan
135	Laki-laki	35	Belum Menikah	Pegawai Swasta	5700000	0	27708600	2800000	11	2804000	Diragukan
136	Laki-laki	51	Menikah	Petani	5500000	3	29086000	3000000	23	1696000	Diragukan
137	Laki-laki	28	Menikah	Mekanik bengkel	5500000	1	21970000	2000000	35	964000	Diragukan
138	Perempuan	44	Menikah	Pengusaha	5000000	3	17194000	1800000	35	785000	Diragukan
139	Laki-laki	28	Belum Menikah	Honorer	5000000	0	28678500	10900000	17	1503000	Diragukan
140	Laki-laki	37	Menikah	Pedagang	5000000	3	18078000	1900000	35	834000	Diragukan
141	Perempuan	38	Widow	Pedagang	5000000	3	19836000	2000000	35	884000	Diragukan
142	Laki-laki	40	Menikah	Pengusaha	5000000	1	29164000	4000000	27	1392000	Diragukan

Repository Universitas Brawijaya
REPOSITORY.UB.AC.ID

No		A2	A3	A4	A5	A6	A7	A8	A9	A10	Klasifikasi
143	Laki-laki	43	Menikah	Wartawan	5000000	6	23335000	4000000	32	964000	Diragukan
144	Laki-laki	49	Menikah	Karyawan swasta	5000000	3	18072000	1900000	17	1300800	Diragukan
145	Laki-laki	54	Menikah	Pengusaha	5000000	4	29080000	3000000	35	1290300	Diragukan
146	Laki-laki	45	Menikah	PNS	5000000	6	18555000	1900000	29	900000	Diragukan
147	Laki-laki	30	Menikah	PNS	5000000	0	21607000	2200000	32	947000	Diragukan
148	Laki-laki	48	Menikah	Buruh	4500000	2	29086000	5000000	35	602000	Diragukan
149	Laki-laki	32	Belum Menikah	Honorer	4500000	0	17921000	2800000	23	1000000	Diragukan
150	Laki-laki	33	Menikah	Pegawai Swasta	4200000	2	27708600	4400000	35	1102000	Diragukan
151	Perempuan	40	Menikah	Penjahit	4000000	1	21040000	2200000	29	1058000	Diragukan
152	Laki-laki	32	Menikah	Pedagang	4000000	4	28678500	2900000	23	1676000	Diragukan
153	Perempuan	38	Menikah	Pedagang	4000000	2	19736000	2000000	35	879000	Diragukan
154	Laki-laki	37	Menikah	Kayawan Kontrak	4000000	2	21040000	2200000	22	612000	Diragukan
155	Perempuan	53	Widow	Penjahit	4000000	2	19629000	2000000	35	863000	Diragukan
156	Laki-laki	34	Menikah	Wiraswasta	4000000	3	29710600	3000000	12	2121000	Diragukan
157	Laki-laki	32	Belum Menikah	Wiraswasta	4000000	0	18991000	1900000	22	1076000	Diragukan
158	Laki-laki	40	Menikah	Wiraswasta	4000000	2	18821000	1900000	29	914000	Diragukan
159	Laki-laki	64	Menikah	Wiraswasta	4000000	3	16544000	3350000	20	916700	Diragukan
160	Laki-laki	24	Belum Menikah	Karyawan Swasta	3500000	0	29086000	3000000	35	1294000	Diragukan
161	Perempuan	35	Menikah	PNS	3500000	3	17806000	2750000	29	838000	Diragukan
162	Perempuan	45	Widow	Wiraswasta	3500000	0	19091000	2200000	29	916000	Diragukan
163	Perempuan	24	Belum Menikah	Pegawai Swasta	3500000	0	18035000	3700000	30	744700	Diragukan
164	Laki-laki	40	Menikah	Pegawai Swasta	3500000	2	18943000	1900000	35	815000	Diragukan
165	Laki-laki	30	Belum Menikah	Pegawai Swasta	3300000	0	16802000	1700000	35	771000	Diragukan
166	Laki-laki	28	Belum Menikah	Pegawai Swasta	3250000	1	33555200	5250000	29	1573300	Diragukan

Repository Universitas Brawijaya
UNIVERSITAS BRAWIJAYA
REPOSITORY.UB.AC.ID

No	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	Klasifikasi
167	Perempuan	34	Menikah	Pegawai Swasta	3100000	0	28308600	4100000	23	1548000	Diragukan
168	Perempuan	32	Belum Menikah	Pegawai Swasta	3100000	0	19300000	2000000	30	868000	Diragukan
169	Laki-laki	27	Belum Menikah	Honorer	3000000	0	28678500	12100000	35	865000	Diragukan
170	Perempuan	25	Belum Menikah	Pedagang	3000000	0	29086000	3000000	35	647000	Diragukan
171	Laki-laki	53	Menikah	Pedagang	3000000	4	19730000	2000000	35	875500	Diragukan
172	Perempuan	42	Menikah	Pegawai Swasta	3000000	3	17436300	5000000	30	651500	Diragukan
173	Perempuan	54	Menikah	PNS	3000000	3	17247000	4250000	32	659500	Diragukan
174	Perempuan	36	Widow	Pegawai Swasta	3000000	0	17436300	3000000	32	730000	Diragukan
175	Perempuan	30	Belum Menikah	Pegawai Swasta	3000000	0	28716000	2900000	29	1452000	Diragukan
176	Perempuan	24	Menikah	Pegawai Swasta	3000000	0	19093000	2000000	30	838000	Diragukan
177	Perempuan	26	Menikah	Pegawai Swasta	2700000	2	17358000	1800000	34	793000	Diragukan
178	Laki-laki	29	Menikah	Pegawai Swasta	2700000	1	16802000	1700000	32	771000	Diragukan
179	Laki-laki	20	Belum Menikah	Pegawai Swasta	2500000	0	16551000	1700000	35	750000	Diragukan
180	Perempuan	44	Menikah	Pegawai Swasta	10000000	0	16451000	3300000	18	1020200	Macet
181	Perempuan	51	Menikah	PNS	10000000	0	17340700	1800000	23	997000	Macet
182	Laki-laki	44	Menikah	PNS	8000000	4	27460500	5600000	35	1055000	Macet
183	Laki-laki	52	Belum Menikah	Pegawai Swasta	7000000	0	16901000	1700000	30	747000	Macet
184	Laki-laki	23	Menikah	Wiraswasta	6000000	2	17101000	1800000	35	762000	Macet
185	Laki-laki	49	Menikah	Wiraswasta	6000000	4	16051000	5000000	27	649600	Macet
186	Laki-laki	41	Menikah	PNS	6000000	3	16051000	2500000	26	782400	Macet
187	Laki-laki	40	Menikah	Wiraswasta	5000000	2	18497000	5000000	30	673600	Macet
188	Laki-laki	58	Menikah	Wiraswasta	5000000	5	16806700	2600000	30	753700	Macet
189	Perempuan	37	Menikah	Wiraswasta	5000000	3	18135000	2750000	26	883800	Macet
190	Laki-laki	39	Menikah	Wiraswasta	5000000	0	17436300	3500000	24	839300	Macet

Repository Universitas Brawijaya
REPOSITORY.UB.AC.ID
UNIVERSITAS BRAWIJAYA

No	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	Klasifikasi
191	Perempuan	32	Menikah	Wiraswasta	4500000	5	17574200	2650000	29	820100	Macet
192	Perempuan	49	Widow	Wiraswasta	4500000	1	16647000	1700000	35	762400	Macet
193	Laki-laki	35	Menikah	Wiraswasta	4500000	4	16019000	3250000	29	705600	Macet
194	Laki-laki	38	Widow	Karyawan swasta	4500000	0	16108500	1650000	29	801000	Macet
195	Laki-laki	41	Menikah	Wiraswasta	4500000	3	18471000	1850000	29	938000	Macet
196	Perempuan	58	Menikah	Ibu Rumah Tangga	4500000	0	34555200	7000000	32	1382300	Macet
197	Laki-laki	47	Menikah	PNS	4500000	2	18135000	2750000	30	813700	Macet
198	Laki-laki	35	Menikah	PNS	4000000	3	16287000	3300000	27	725900	Macet
199	Laki-laki	32	Menikah	Wiraswasta	4000000	2	17806000	2700000	17	1224100	Macet
200	Laki-laki	47	Menikah	Wiraswasta	4000000	1	16313000	1025000	35	673500	Macet
201	Laki-laki	34	Menikah	Petani	4000000	3	17221000	2000000	29	827000	Macet
202	Laki-laki	49	Menikah	PNS	4000000	3	16451000	2500000	30	740700	Macet
203	Laki-laki	52	Menikah	Wiraswasta	4000000	3	26440600	5500000	32	1058100	Macet
204	Perempuan	40	Menikah	PNS	4000000	1	16806700	3400000	30	698300	Macet
205	Laki-laki	58	Menikah	Wiraswasta	4000000	4	16451000	3300000	11	1559300	Macet
206	Laki-laki	31	Belum Menikah	Pegawai Swasta	4000000	0	33805200	5400000	29	1512000	Macet
207	Laki-laki	43	Menikah	Pegawai Swasta	4000000	0	16901000	1700000	35	747000	Macet
208	Perempuan	33	Menikah	Pegawai Swasta	3800000	0	18035000	3250000	26	826700	Macet
209	Perempuan	33	Menikah	Pegawai Swasta	3800000	0	18035000	3250000	24	1057800	Macet
210	Laki-laki	34	Menikah	Wiraswasta	3750000	3	18471000	2800000	35	790600	Macet
211	Perempuan	27	Belum Menikah	Wiraswasta	3500000	0	17321000	1800000	30	762000	Macet
212	Perempuan	35	Menikah	Pegawai Swasta	3500000	3	19093000	2000000	27	953000	Macet
213	Perempuan	33	Menikah	PNS	3500000	2	16051000	2750000	30	707900	Macet
214	Perempuan	52	Menikah	Buruh	3300000	6	14259000	1125000	23	788000	Macet

Repository Universitas Brawijaya
REPOSITORY.UB.AC.ID
UNIVERSITAS BRAWIJAYA

No	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	Klasifikasi
215	Perempuan	46	Belum Menikah	Wiraswasta	3200000	0	16437000	1750000	35	747700	Macet
216	Laki-laki	42	Menikah	Pegawai Swasta	3200000	1	16901000	2700000	26	793000	Macet
217	Laki-laki	42	Menikah	Wiraswasta	3000000	2	16287000	2450000	35	697500	Macet
218	Laki-laki	30	Menikah	Wiraswasta	3000000	1	18071000	2750000	29	858900	Macet
219	Perempuan	29	Belum Menikah	Pegawai Bank	3000000	0	13204000	725000	29	626100	Macet
220	Perempuan	34	Menikah	Wiraswasta	3000000	3	16394000	2275000	35	636300	Macet
221	Perempuan	30	Menikah	Pegawai Swasta	3000000	1	14259000	2575000	29	594700	Macet
222	Laki-laki	32	Belum Menikah	Wiraswasta	3000000	0	17695000	3550000	35	690300	Macet
223	Perempuan	34	Menikah	Wiraswasta	3000000	2	21354000	2150000	35	910900	Macet
224	Laki-laki	55	Belum Menikah	Supir	3000000	0	17436300	2650000	24	904000	Macet
225	Perempuan	41	Menikah	Wiraswasta	3000000	0	16451000	2500000	26	804500	Macet
226	Laki-laki	23	Belum Menikah	Pegawai Swasta	3000000	0	17340700	1800000	35	763000	Macet
227	Laki-laki	34	Belum Menikah	PNS	2800000	0	19285000	2900000	26	939100	Macet
228	Perempuan	26	Belum Menikah	Pegawai Swasta	2800000	0	16051000	2500000	24	831200	Macet
229	Laki-laki	37	Menikah	Pegawai Swasta	2800000	2	31239000	15000000	23	1122000	Macet
230	Laki-laki	25	Belum Menikah	Pedagang	2500000	0	16806700	3500000	30	693500	Macet
231	Laki-laki	27	Menikah	Wiraswasta	2500000	1	18135000	3250000	26	856600	Macet
232	Perempuan	35	Menikah	Pegawai Swasta	2500000	0	17436300	2700000	26	848000	Macet
233	Perempuan	32	Menikah	Pegawai Swasta	2500000	1	18135000	2750000	14	1511100	Macet
234	Laki-laki	25	Belum Menikah	Wiraswasta	2300000	0	15921000	875000	23	872300	Macet
235	Laki-laki	35	Belum Menikah	Wiraswasta	2000000	0	16394000	1325000	29	746500	Macet
236	Laki-laki	46	Menikah	Wiraswasta	2000000	3	13204000	725000	35	561700	Macet
237	Laki-laki	39	Menikah	Wiraswasta	1800000	2	15336000	775000	35	622800	Macet
238	Laki-laki	33	Menikah	Wiraswasta	1800000	1	15336000	775000	35	622800	Macet

Repository Universitas Brawijaya

No	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	Klasifikasi
239	Perempuan	30	Menikah	Pegawai Swasta	1500000	1	14259000	1025000	35	602400	Macet
240	Laki-laki	37	Menikah	Pegawai Swasta	1350000	2	13204000	725000	35	566700	Macet

Keterangan:

A1= Jenis kelamin
A2= Usia
A3= Status
A4= Pekerjaan
A5= Pendapatan
A6= Tanggungan
A7= Harga On The Road (OTR)
A8= DP
A9= Tenor
A10= Angsuran
Klasifikasi= label kelas kredit kendaraan

Repository Universitas Brawijaya
UNIVERSITAS BRAWIJAYA
REPOSITORY.UB.AC.ID