

# DATA MINING UNTUK KLASIFIKASI PINJAMAN KREDIT PENSIUNAN MENGUNAKAN ALGORITMA K-NEAREST NEIGHBOR (STUDI KASUS: PT. BANK BUKOPIN Tbk)

Vidya Alfani

Prodi Teknik Informatika STMIK Budi Darma, Medan, Indonesia  
Jl. Sisingamangaraja No. 338 Simpang Limun, Medan

## ABSTRAK

Kredit Pensiunan merupakan kredit yang sifatnya konsumtif, yang hanya diberikan kepada para pensiunan, yang berasal dari lingkungan pensiunan karyawan (PNS, BUMN/BUMD) yang didasari oleh perjanjian kerjasama antara pihak Bank oleh pengelola Dana Pensiun. Permasalahan yang sering timbul adalah terdapatnya perbedaan nominal gaji pensiunan berdasarkan usia dan golongan terakhir dari berbagai instansi sehingga proses penentuan nilai plafond memerlukan proses klasifikasi sesuai data yang diajukan calon debitur untuk memudahkan marketing dalam menyampaikan jumlah pinjaman yang bisa diberikan. Data mining dapat diartikan sebagai pengestrakan informasi baru yang diambil dari bongkahan data besar yang membantu dalam pengambilan keputusan. K-Nearest Neighbor merupakan sebuah metode untuk klasifikasi terhadap objek berdasarkan data pembelanjaraan yang jaraknya paling dekat dengan objek tersebut. Teknik ini sangat sederhana dan mudah diimplementasikan. Dengan penggunaan tools weka, hasil dari proses penerapan Algoritma K-Nearest Neighbor adalah informasi atau knowledge dimana informasi tersebut dapat menjadi alternatif yang dapat digunakan untuk mengklasifikasi besar plafond pinjaman kredit pensiunan.

**Kata Kunci:** Kredit Pensiunan, Data Mining, K-Nearest Neighbor, Weka

## I. PENDAHULUAN

Kredit Pensiunan merupakan kredit yang sifatnya *konsumtif*, yang hanya diberikan kepada para pensiunan, yang berasal dari lingkungan pensiunan karyawan (PNS, BUMN/BUMD) yang didasari oleh perjanjian kerjasama antara pihak Bank oleh pengelola Dana Pensiun. PT. Bank Bukopin Tbk adalah salah satu lembaga keuangan yang menyediakan beberapa produk yang ditawarkan kepada para pensiunan seperti tabungan pensiunan dan kredit pensiun. Sementara bentuk layanan bank bisa berupa jasa pembayaran pensiun, jasa *transfer*, jasa kliring dan pengelolaan program daya. Salah satu program yang paling diminati adalah produk kredit pensiunan. Permasalahan yang sering timbul adalah terdapatnya perbedaan nominal gaji pensiunan berdasarkan usia dan golongan terakhir dari berbagai instansi sehingga proses penentuan nilai *plafond* memerlukan proses klasifikasi sesuai data yang diajukan calon *debitur* untuk memudahkan *marketing* dalam menyampaikan jumlah pinjaman yang bisa diberikan.

*Data mining* adalah serangkaian proses untuk menggali nilai tambah dari suatu kumpulan data berupa pengetahuan yang selama ini tidak diketahui secara manual. *Data mining* juga dapat diartikan sebagai pengestrakan informasi baru yang diambil dari bongkahan data besar yang membantu dalam pengambilan keputusan [1]. Banyak algoritma dalam metode klasifikasi *data mining* salah satunya adalah Algoritma *K-Nearest Neighbor*. *K-Nearest Neighbor* merupakan sebuah metode untuk klasifikasi terhadap

objek berdasarkan data pembelanjaraan yang jaraknya paling dekat dengan objek tersebut. Teknik ini sangat sederhana dan mudah diimplementasikan. Berdasarkan penelitian sebelumnya yang dilakukan oleh, Nursalim Suprapedi dan H.Himawan menyatakan bahwa algoritma *k-nearest neighbor* memiliki kinerja terbaik untuk klasifikasi bidang kerja kelulusan dengan nilai akurasi [2].

## II. TEORITIS

### A. Data Mining

Menurut Connolly dan Begg, *Data mining* adalah suatu proses ekstraksi atau penggalian data yang belum diketahui sebelumnya, namun dapat dipahami dan berguna dari *database* yang besar serta digunakan untuk membuat suatu keputusan bisnis yang sangat penting [3]. *Data mining* adalah salah satu langkah dalam proses KDD secara keseluruhan. Secara umum, *data mining* digunakan oleh banyak peneliti sebagai sinonim dari proses KDD. *Knowledge Discovery in Database* berkaitan dengan proses penemuan pengetahuan yang diterapkan pada *database*. [3]. Proses KDD dapat dilihat pada gambar 2.1 yang terdiri dari tahapan-tahapan sebagai berikut [4]:

### B. Algoritma K- Nearest Neighbor (KNN)

Algoritma *Nearest Neighbor* (kadang disebut Algoritma *K- Nearest Neighbor*/K-NN) merupakan algoritma yang melakukan klasifikasi berdasarkan

kedekatan lokasi (jarak) suatu data dengan data yang lain [1]. *Nearest Neighbor* adalah suatu pendekatan untuk menghitung kedekatan antara kasus baru dengan kasus lama, yaitu berdasarkan pada pencocokan bobot dari sejumlah fitur yang ada [5]. Rumus untuk melakukan penghitungan kedekatan antara kedua dua kasus adalah sebagai berikut:

$$\text{Similarity}(T, S) = \frac{\sum_{i=1}^n f(T_i, S_i) * W_i}{w_i} \dots\dots\dots 1$$

Dimana:

- T : kasus baru
- S : kasus dalam penyimpanan
- N : jumlah atribut tiap kasus
- i : atribut individu 1 s/d n
- f : fungsi kesamaan atribut i antara kasus T & S
- w : bobot pada atribut yang ke i

Adapun urutan Algoritma KNN sebagai berikut [5]:

1. Menentukan parameter K (jumlah tetangga paling dekat)
2. Menghitung kuadrat jarak euclidian (*euclidean distance*) masing-masing obyek terhadap data *sampel* yang diberikan  

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \dots\dots\dots 2$$
3. Mengurutkan objek-objek tersebut ke dalam kelompok yang mempunyai jarak *euclid* terkecil
4. Mengumpulkan kategori Y (klasifikasi *nearest neighbor*)
5. Dengan menggunakan kategori mayoritas, maka dapat diprediksikan nilai *query instance* yang telah dihitung.

**C. Klasifikasi**

Klasifikasi pertama kali diterapkan pada bidang tanaman yang mengklasifikasi suatu spesies tertentu, seperti yang dilakukan oleh Carolus von Linne (atau dikenal dengan nama Carolus Linnaeus) yang pertama kali mengklasifikasikan spesies berdasarkan karakteristik fisik. Selanjutnya dia dikenal sebagai bapak klasifikasi [6].

**D. Kredit Pensiun**

Kredit adalah kemampuan untuk melaksanakan suatu pembelian atau mengadakan suatu pinjaman dengan suatu janji pembayarannya akan dilakukan dengan ditangguhkan pada suatu jangka waktu tertentu yang telah disepakati (IKAPI). Sedangkan pengertian yang lebih mapan untuk kegiatan perbankan di Indonesia juga dirumuskan dalam bab I, pasal 1,2 undang-undang Pokok Perbankan No. 14 Tahun 1967 yang merumuskan : “Kredit adalah penyediaan uang atau tagihan yang dapat disamakan dengan itu, berdasarkan persetujuan pinjam-meminjam antara bank dengan pihak lain

dalam hal dimana peminjam berkewajiban melunasi hutangnya setelah jangka waktu tertentu dengan jumlah bunga yang telah ditentukan (IKAPI) [7].

Pensiun adalah penghasilan yang diterima setiap bulan oleh seorang bekas pegawai yang tidak dapat bekerja lagi untuk membiayai kehidupan selanjutnya agar tidak terlantar apabila tidak berdaya lagi untuk mencari penghasilan yang lain. Berdasarkan UU No.11 Tahun 1969, “Pensiun diberikan sebagai jaminan hari tua dan sebagai penghargaan atas jasa-jasa pegawai negeri selama bertahun-tahun bekerja dalam dinas pemerintah” [8].

**III. ANALISA**

**A. Analisa Masalah**

Dengan studi kasus pada PT. Bank Bukopin, Tbk untuk mengklasifikasi pinjaman kredit pensiun sebelumnya. Khususnya data calon nasabah pensiun dimulai dari 1 tahun pada tahun 2016, dalam pembahasan ini data yang diambil hanya menggunakan data nasabah bagi pensiun pegawai negeri sipil (PNS) dan data yang diambil hanya menggunakan data 1 bulan sebagai *sampel*. Berikut ini adalah data buku bulanan yang akan dijadikan sampel untuk analisa dan untuk pengujian diambil bulan Desember 2016.

Tabel 1 Data Nasabah Pensiunan PNS (Tranformasi Data)

No	Nama Nasabah	Usia	Gol	Gaji	Pinjaman
1	Fauzah S	63-75 thn	4A	3.5-6 juta	230 juta
2	Sitor S.	63-75 thn	4A	3-3.4 juta	175 juta
3	Mutia D.	60-62 thn	4A	3-3.4 juta	175 juta
4	Yusni Aida	60-62 thn	4C	3.5-6 juta	230 juta
5	Juliper P.	60-62 thn	4A	2- 2.9 juta	100 juta
6	Bertelena S.	60-62 thn	4A	2- 2.9 juta	100 juta
7	Tiarma Butar-Butar	60-62 thn	4A	2- 2.9 juta	100 juta
8	Lamsahat Silitonga	60-62 thn	3D	1-1.9 juta	60 juta
9	Emseri Marbun	60-62 thn	4A	2- 2.9 juta	60 juta
10	Nurleli Pane	60-62 thn	3B	2- 2.9 juta	100 juta
11	Halimah Tussakdi h M.	60-62 thn	2A	1-1.9 juta	60 juta
12	Suhasni	60-62 thn	3A	2- 2.9 juta	100 juta
13	Sulasiah	60-62 thn	3D	3-3.4 juta	175 juta

No	Nama Nasabah	Usia	Gol	Gaji	Pinjaman
14	Kasman Sirait	60-62 thn	4A	3-3.4 juta	175 juta
15	Suarno	63-75 thn	4A	3.5-6 juta	230 juta
16	Wilson Mulatu S.	63-75 thn	4A	3-3.4 juta	175 juta
17	M. Nasib Suherman	63-75 thn	4B	3.5-6 juta	230 juta
18	Bebas Damanik SE	63-75 thn	4A	3.5-6 juta	230 juta
19	Sunarno	60-62 thn	4A	3-3.4 juta	175 juta
20	Bakhtiar H.	63-75 thn	3D	2- 2.9 juta	100 juta
21	Lebanus Str	63-75 thn	4A	2- 2.9 juta	100 juta
22	Martha Sinuraya	63-75 thn	4A	3-3.4 juta	175 juta
No	Nama Nasabah	Usia	Gol	Gaji	Pinjaman
23	Farina Wati Br Purba	63-75 thn	3D	2- 2.9 juta	100 juta
24	Nurmiah	60-62 thn	4A	3-3.4 juta	175 juta
25	Benti G	60-62 thn	4A	3-3.4 juta	175 juta

Keterangan:

Untuk atribut tujuan Pinjaman dikelompokkan menjadi 8 kelompok, yaitu:

- A = Max 230 juta / 15 thn
- B = Max 230 juta / 10 thn
- C = Max 175 juta/ 15 thn
- D = Max 175 juta/ 10 thn
- E = Max 100 juta/ 15 thn
- F = Max 100 juta/ 10 thn
- G = Max 60 juta/ 15 thn
- H = Max 60 juta/ 10 thn

### B. Penyelesaian Algoritma KNN

Setelah dilakukan transformasi data, maka dilakukan langkah-langkah berikut:

#### 1. Penentuan Nilai Pembobotan

Untuk melakukan pengukuran jarak antara atribut, maka akan dilakukan pembobotan pada atribut.

Tabel 2 Pembobotan Atribut

Atribut	Bobot
Usia	0.75
Gol	0.25
Gaji	1

#### 2. Penentuan Nilai Kedekatan Atribut

Untuk penyelesaian kasus pada fase awalnya, kita perlu untuk menentukan nilai kedekatan antara setiap nilai-nilai atribut. Berikut ini adalah kedekatan nilai-nilai dari setiap atribut kondisinya.

Tabel 3 Kedekatan Nilai-Nilai Atribut

Atribut	Nilai Atribut 1	Nilai Atribut 2	Bobot
Lahir	63-75 thn	63-75 thn	0.75
	60-62 thn	60-62 thn	0.75
	63-75 thn	60-62 thn	0.50
	60-62 thn	63-75 thn	0.50
Gol	2A	2A	0.25
	2A	3A	0.20
	2A	3B	0.20
	2A	3D	0.20
	2A	4A	0.20
	2A	4B	0.20
	2A	4C	0.20
	3A	3A	0.25
	3A	3B	0.20
	3A	3D	0.20
	3A	4A	0.20
	3A	4B	0.20
	3A	4C	0.20
	3B	3B	0.25
	3B	3D	0.20
	3B	4A	0.20
	3B	4B	0.20
	3B	4C	0.20
	3D	3D	0.20
	3D	4A	0.20
	3D	4B	0.20
	3D	4C	0.20
	4A	4A	0.25
	4A	4B	0.20
4A	4C	0.20	
4B	4B	0.20	
4B	4C	0.20	
4C	4C	0.25	
Gaji	1-1.9 juta	1-1.9 juta	1
	1-1.9 juta	2- 2.9 juta	0.75
	1-1.9 juta	3-3.4 juta	0.75
	1-1.9 juta	3.5-6 juta	0.75
	2- 2.9 juta	2- 2.9 juta	1
	2- 2.9 juta	3-3.4 juta	0.75
Gaji	2- 2.9 juta	3.5-6 juta	0.75
	3-3.4 juta	3-3.4 juta	1
	3-3.4 juta	3.5-6 juta	0.75
	3.5-6 juta	3.5-6 juta	1

#### 3. Menghitung Kedekatan Atribut Kasus Baru (*Data Testing*) Untuk mengetahui hasil analisis dari algoritma *k-nearest neighbor* maka dilakukan perhitungan manual dengan menggunakan data training (kasus lama) dan data testing (kasus baru).

Tabel 4 Data Calon Nasabah Pensiun Baru

Nama	Lahir	Gol	Gaji
Sutraswo	63 Tahun	4A	3 juta

Untuk menghitung kedekatan kasus antara data training dan data testing di atas, berikut ini adalah algoritma penyelesaiannya.

1. Menghitung Nilai Kedekatan Atribut Kondisi Kasus baru dengan Kasus No 1

Keterangan :

- a. Kedekatan Kasus Baru Dengan Kasus 1

Nama	Lahir	Gol	Gaji
Sutraswo	63-75 thn	4A	3-3.4 juta
Fauzah S	63-75 thn	4A	3.5-6 juta
Nilai Kedekatan	0.75	0.25	0.75
Nilai Atribut	a	c	e

- b. Bobot Kasus 1

Atribut	Bobot	Nilai Atribut
Lahir	0.75	b
Gol	0.25	d
Gaji	1	f

Hitung :

$$Similarity(T, S) = \frac{\sum_{i=1}^n f(T_i, S_i) * W_i}{W_i}$$

$$Jarak = \frac{(a*b) + (c*d) + (e*f)}{b + d + f}$$

$$Jarak = \frac{(0.75*0.75) + (0.25*0.25) + (0.75*1)}{0.75 + 0.25 + 1}$$

$$Jarak = 1.375 / 2 = 0.6875$$

Lakukan dengan cara sama sampai kasus No. 25

Setelah dilakukan penghitungan nilai kedekatan kasus baru dengan seluruh kasus lama dengan Algoritma *K-Nearest Neighbor*, maka berikut adalah hasil rekapitulasi nilai kedekatan kasus baru dengan kasus lama:

Tabel 5 Rekapitulasi Nilai Kedekatan

No	Kasus	Nilai Kedekatan
1	Kasus 1	0.6875
2	Kasus 2	<b>0.8125</b>
3	Kasus 3	0.71875
4	Kasus 4	0.587
5	Kasus 5	0.59375
6	Kasus 6	0.59375
7	Kasus 7	0.59375
8	Kasus 8	0.587
9	Kasus 9	0.59375
10	Kasus 10	0.587
11	Kasus 11	0.587
12	Kasus 12	0.587
13	Kasus 13	0.7125
14	Kasus 14	0.718
15	Kasus 15	0.6875

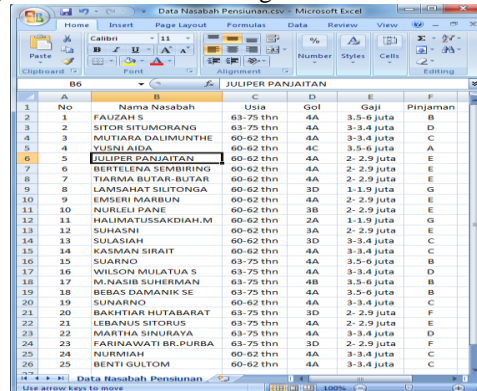
No	Kasus	Nilai Kedekatan
16	Kasus 16	<b>0.8125</b>
17	Kasus 17	0.5875
18	Kasus 18	0.6875
19	Kasus 19	0.71875
20	Kasus 20	0.5875
21	Kasus 21	0.59375
22	Kasus 22	<b>0.8125</b>
23	Kasus 23	0.5875
24	Kasus 24	0.71875
25	Kasus 25	0.71875

Berdasarkan tabel diatas untuk nilai Kedekatan tertinggi adalah pada nilai Kedekatan dengan Kasus 2, Kasus 16, dan Kasus 21 dengan nilai kedekatan 0.8125. Jadi, untuk soal di atas maka nasabah atas nama "Sutraswo" pada nilai atribut keterangannya bernilai "D" atau yang mendapatkan pinjaman Max 175 juta/ 10 tahun.

#### IV. IMPLEMENTASI

##### A. Implementasi Sistem

Implementasi sistem program ini mencakup spesifikasi kebutuhan perangkat keras (*hardware*) dan spesifikasi perangkat lunak (*software*). Tampilan program terdiri dari *print screen* dari tampilan *input*, *output*, dan proses percobaan. Adapun tampilan program WEKA adalah sebagai berikut:

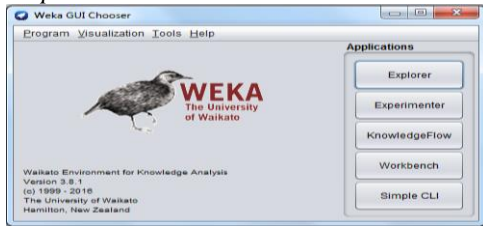


Gambar 1 Tampilan *Input* Data Pensiunan Pegawai PNS

Data nasabah pensiunan PNS yang *diinput* dari *database* telah selesai. Maka tahap selanjutnya adalah proses data menggunakan aplikasi WEKA sebagai berikut:

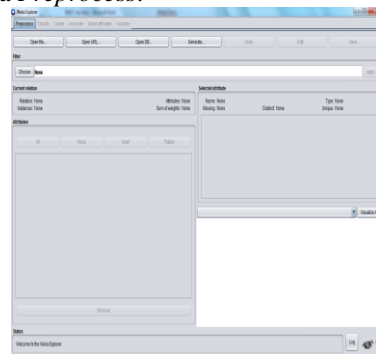
1. Buka program WEKA yang telah *diinstal*, setelah proses maka akan muncul tampilan awal aplikasi WEKA seperti pada gambar 3 Pada halaman awal terdapat beberapa *fitur*, salah satunya adalah *explorer*. *Explorer* berguna untuk memulai proses

analisis baru. Untuk memulai proses baru kemudian klik *explorer*.



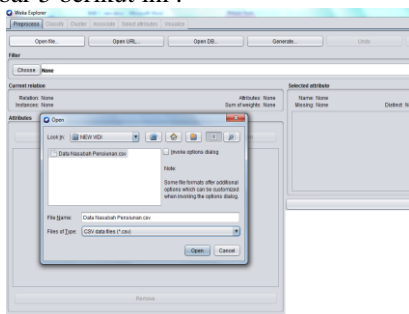
Gambar 2 Tampilan Awal WEKA

2. Kemudian akan muncul tampilan yang terlihat pada gambar 4 dan untuk melakukan proses data, langkah selanjutnya adalah memasukkan data nasabah pensiunan yang sudah tersimpan dalam format data \*.CSV kedalam *microsoft excel*, untuk memanggil data nasabah pensiunan yaitu dengan cara klik *open file* pada *Preprocess*.



Gambar 3 Tampilan Preprocess WEKA

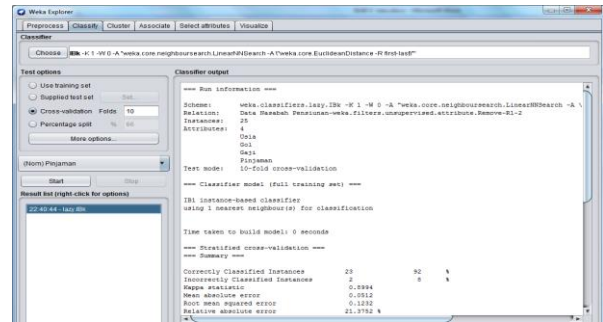
3. Setelah mengklik *open file* maka sistem akan menampilkan yaitu mencari dimana data yang akan di *import* yaitu data yang sudah di simpan pada format \*.CSV atau lebih jelas dapat dilihat pada gambar 5 berikut ini :



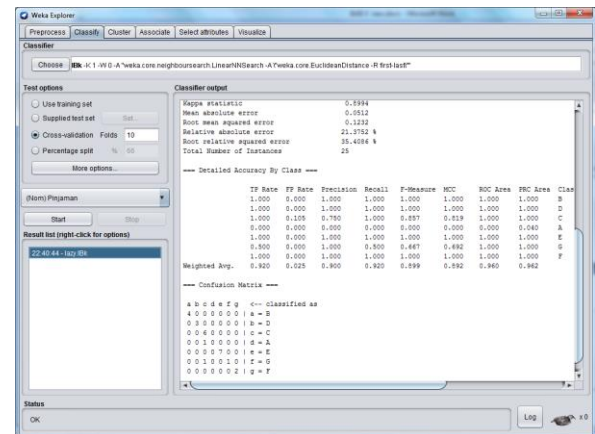
Gambar 4 Tampilan Pencarian Data

## B. Hasil Pengujian Program

Hasil akhir atau *output* dari aplikasi WEKA dengan menggunakan algoritma *nearest neighbour* dapat dilihat pada gambar 6 berikut:



Gambar 5 Tampilan Output Data 1



Gambar 6 Tampilan Output Data 2

## V. KESIMPULAN

Berdasarkan uraian yang telah dibuat tentang implementasi Algoritma *K-Nearest Neighbor* dalam penentuan klasifikasi pinjaman pensiunan, maka dapat disimpulkan bahwa:

1. Prosedur awal bagi debitur baru harus mengikuti perlengkapan berkas seperti SK Pensiun, Karip, dan Sebagainya yang telah ditetapkan Bank Bukopin.
2. Algoritma *K-Nearest Neighbor* dapat digunakan dalam Pengklasifikasian jumlah plafond pinjaman kredit Pensiunan pada PT.Bank Bukopin Tbk.
3. Pengujian penentuan Plafond menggunakan *weka* dimulai dengan penginputan data nasabah menjadi database pada *Ms.Excel* dengan ekstensi CSV. Mengelompokkan data dengan algoritma *K-Nearst Neighbor* dilakukan dengan cara menentukan atribut apa saja yang di dapat dari hasil *cleaning*, sehingga dapat menghasilkan jarak tetangga terdekatdari masing-masing atribut.

## REFERENCES

[1] Prasetyo, *Data Mining- Konsep dan Aplikasi Menggunakan Matlab*. Yogyakarta: C.V ANDI OFFSET (Penerbit Andi), 2012.

- [2] Nursalim et al, "Klasifikasi Bidang Kerja Kelulusan Menggunakan Algoritma K-Nearest Neighbor," *Jurnal Teknologi Informasi*, vol. 10, no. 1, April 2014.
- [3] K.J. Arta et al, "Data Mining Rekomendasi Calon Mahasiswa Berprestasi Di STMIK Denpasar Menggunakan Metode Technique For Others Reference By Similary To Ideal Solution," *Jurnal Sains dan Teknologi*, vol. 5, no. 2, Oktober 2016
- [4] J. Han et al. (2012) *Data Mining: Concepts and Techniques Third Edition*.
- [5] Kusriani & A. T. Luthfi, *Algoritma Data Mining*. Yogyakarta: C.V ANDI OFFSET (Penerbit Andi), 2009.
- [6] P. P. Widodo et al, *Penerapan Data Mining Dengan Matla*. Bandung: Penerbit Rekayasa Sains, 2013.
- [7] M. B. Suwadnyaa et al, "Fuzzy Inference Sistem Mamdani Untuk Penentuan Kredit Pada KPN Estika Dewata," *Jurnal Nasional Pendidikan Teknik Informatika (JANAPATI)*, vol. 1, no. 2, Jul 2013.
- [8] J. R. Uhise, "Analisis Penerapan Sistem Pengendalian Manajemen Penyaluran Kredit Pada BRI Kota Manado," *Jurnal EMBA*, vol. 1, no. 3, September 2013.
- [9] F. L. Simon & J. Santana, *Data Mining : Meramalkan Bisnis Perusahaan*. Jakarta: PT. Elex Media Komputind, 2010.
- [10] J. Arifin & A. Fauzi, *Mengupas Tuntas Microsoft Office Excel 2007*. Jakarta: PT. Elex Media Komputindo, 2007.