

# PENERAPAN DATA *MINING* DENGAN METODE *INTERPOLASI* UNTUK ESTIMASI PREMI BERDASARKAN PENDAPATAN NASABAH ASURANSI PRUDENTIAL DI YOGYAKARTA

**Pandapotan, Anton Setiawan Honggowibowo, Dwi Nugraheny**  
Teknik Informatika STTA Yogyakarta  
informatika@stta.ac.id

## *Abstract*

*Premi in a definition can be interpreted as a fee to be paid by the customer, which is a combination of the overall cost of insurance benefits that used, and sometimes also include the amount of money invested by customers.*

*The output of the data mining is a customers is accepted according to the prudential value of the estimated premi be gained from the number of customer revenue.*

*Final project is, created a program of web-based application system to estimate the premi to make it easier for admin (prudential employee) to determine the amount of premi to be paid by the customer. In this program, admin (prudential employee) just need to know the amount of customer's income that the program will then process the results of the estimated premi.*

*The test results is concluded that the program estimates the premi can be used easily by the admin (prudential employee) and simplify admin decisions premi estimate to the customer.*

***Keywords: Premi, Clustering, Interpolasi, DataMining***

## Abstrak

Premi dapat diartikan sebagai biaya yang harus dibayar oleh nasabah, yang merupakan kombinasi dari biaya keseluruhan manfaat asuransi yang digunakan, dan kadang-kadang juga termasuk jumlah uang yang diinvestasikan oleh nasabah. *Output* dari data *mining* ini adalah calon nasabah diterima sebagai nasabah sesuai dengan hasil nilai estimasi premi yang akan diperoleh dari jumlah pendapatan nasabah.

Pada tugas akhir ini dibuat sebuah program dalam bentuk sistem aplikasi berbasis web estimasi premi untuk memudahkan admin (Pegawai Prudential) dalam menentukan jumlah premi yang harus dibayar oleh nasabah. Dalam program ini admin (Pegawai Prudential) hanya perlu mengetahui pendapatan nasabah yang kemudian program ini akan memproses hasil dari estimasi premi.

Dari hasil pengujian disimpulkan bahwa program ini dapat digunakan dengan mudah oleh admin (Pegawai Prudential) dalam membantu memberikan keputusan estimasi premi kepada nasabah.

***Kata Kunci: Premi, Clustering, Interpolasi, Data Mining***

## 1. Latar Belakang

Prediksi estimasi premi sangat penting bagi suatu perusahaan asuransi. Estimasi premi sendiri tidak hanya didasarkan pada keuntungan yang didapatkan dari asuransi, melainkan juga dari faktor pendapatan. Jika premi konsumen tidak didukung dengan pendapatan yang memadai untuk membayar premi asuransi, tentunya kondisi ini tidak menguntungkan bagi perusahaan. Dengan adanya estimasi premi akan produk asuransi, perusahaan asuransi, dalam hal ini perusahaan Asuransi PT. Prudential Yogyakarta, dapat mengambil suatu keputusan yang strategis untuk para agen-agensya dalam mempromosikan produk asuransinya.

Teknologi data *mining* dapat dijadikan alat bantu dalam mengolah data yang berjumlah besar untuk dijadikan data prediksi. Dengan memanfaatkan data masukan berupa data nasabah, pendapatan nasabah, dan transaksi pembayaran premi asuransi, dapat dibangun aplikasi data *mining* dengan metode interpolasi untuk estimasi premi asuransi.

Keluaran atau *output* dari proses data mining ini adalah masyarakat dapat diterima sebagai nasabah prudential sesuai dengan hasil nilai estimasi premi yang akan diperoleh dari jumlah pendapatan nasabah.

## 2. Data Mining

Kemajuan dalam pengumpulan data dan teknologi penyimpanan yang cepat memungkinkan organisasi menghimpun jumlah data yang sangat besar. Alat dan teknik analisis data yang tradisional tidak lagi dapat digunakan untuk mengekstrak informasi dari data yang sangat besar. Untuk itu diperlukan suatu metoda baru yang dapat menjawab kebutuhan tersebut.

Menurut (Therling K., 2006), data *mining* adalah kombinasi secara logis antara pengetahuan data, dan analisis statistik yang dikembangkan dalam pengetahuan bisnis atau suatu proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, tiruan dan *machine-learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat bagi pengetahuan yang terkait dari berbagai *database* besar.

Pada data mining terdapat 4 tahapan mulai dari seleksi data sampai dengan hasil informasi yang telah di dapat dan dimengerti oleh pihak yang berkepentingan. Adapun penjelasan sebagai berikut.

- a. Tahap 1 adalah *Data Selection* yaitu Menciptakan himpunan data target, pemilihan himpunan data, atau memfokuskan pada subset variabel atau sampel data, tempat penemuan (*discovery*) akan dilakukan.
- b. Tahap 2 adalah *Pre-processing/Cleaning* yaitu : Pemrosesan pendahuluan dan pembersihan data yang merupakan operasi dasar seperti penghapusan *noise*. Sebelum proses data *mining* dapat dilaksanakan, perlu dilakukan proses *cleaning* pada data yang menjadi fokus *Knowledge Discovery in Databases* (KDD).
- c. Tahap 3 adalah *Data mining* yaitu Pemilihan tugas *data mining*; pemilihan *goal* dari proses KDD (*Knowledge Discovery in Databases*) misalnya klasifikasi, *regresi*, *clustering*, dll.
- d. Tahap 4 adalah *Interpretation/ Evaluation* yaitu Penerjemahan pola-pola yang dihasilkan dari data *mining*. Pola informasi yang dihasilkan dari proses data *mining* perlu ditampilkan dalam bentuk yang mudah dimengerti oleh pihak yang berkepentingan.

### Clustering

Clustering adalah proses mengelompokkan obyek berdasarkan informasi yang diperoleh dari data yang menjelaskan hubungan antar obyek dengan prinsip untuk memaksimalkan kesamaan antar anggota satu kelas dan meminimumkan kesamaan antar kelas/cluster. Tujuannya untuk menemukan cluster yang berkualitas dalam waktu yang layak.

Clustering dalam data mining berguna untuk menemukan pola distribusi di dalam sebuah data set yang berguna untuk proses analisa data. Kesamaan obyek biasanya diperoleh dari kedekatan nilai-nilai atribut yang menjelaskan obyek-obyek data, sedangkan obyek-obyek data biasanya direpresentasikan sebagai sebuah titik dalam ruang multidimensi.

### 3. Perancangan Sistem

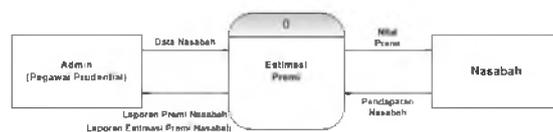
#### Perancangan Diagram Alir Data

Diagram Alir Data merupakan suatu model logika data atau proses yang dibuat untuk menggambarkan dari mana asal data dan ke mana tujuan data yang keluar dari sistem, letak data disimpan, proses apa yang menghasilkan data tersebut dan interaksi antara data yang tersimpan serta proses yang dikenakan pada data tersebut. Dengan DAD ini dapat diketahui letak data disimpan dan bagaimana transformasi datanya.

#### Diagram Konteks

Diagram konteks merupakan diagram yang menjelaskan proses perjalanan data dari satu atau beberapa source untuk mencapai suatu tujuan tertentu yang mana pada proses perjalanan data tersebut hanya terdapat satu proses saja yang digambarkan dalam bentuk umum.

Diagram konteks aplikasi untuk estimasi premi yang dibuat dapat dilihat pada Gambar 1.

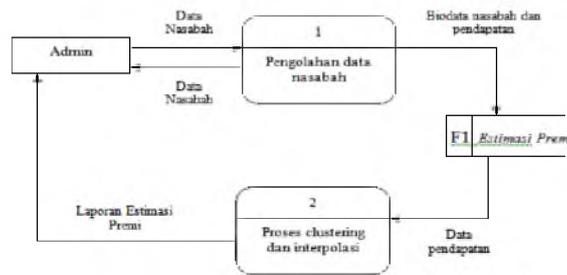


Gambar 1 Diagram Konteks untuk aplikasi estimasi premi

Pada Gambar 1 terdapat 2 entitas, yaitu admin dan nasabah. Admin mengolah Data Nasabah yang kemudian dilakukan proses estimasi premi sehingga diperoleh nilai premi yang diterima untuk nasabah. Selanjutnya nasabah memberikan pendapatan yang kemudian dilakukan estimasi premi sehingga didapatkan laporan premi nasabah yang merupakan laporan estimasi premi bagi nasabah.

#### DAD Level 1

DAD level 1 seperti yang ditampilkan pada gambar 2 merupakan pengembangan dari diagram konteks, dimana pada diagram ini menjelaskan bahwa admin melakukan proses pengolahan data nasabah kemudian dilakukan proses estimasi premi sesuai dengan pendapatan nasabah.



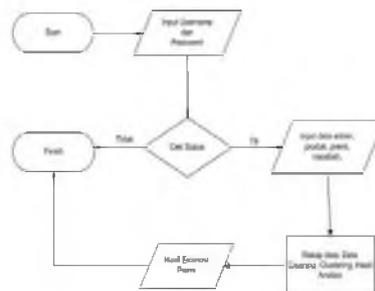
Gambar 2 DAD Level 1 dari Estimasi Premi yang dilakukan oleh admin

Pada proses pengolahan Data Nasabah Gambar 2 diketahui bahwa admin melakukan pengolahan data nasabah yang berupa biodata nasabah, pendapatan yang kemudian nantinya akan disimpan dalam *database* estimasi premi. Selanjutnya estimasi premi database yang disimpan. kemudian data pendapatan nasabah diproses *clustering* dan *interpolasi* dilakukan untuk mengelompokkan data-data pendapatan nasabah kemudian dicari nilai tengahnya dengan cara *interpolasi* sehingga diperoleh nilai estimasi premi untuk nasabah.

### Diagram Alir (Flowchart) Sistem

*Flowchart* adalah penggambaran secara grafik dari langkah-langkah dan urutan prosedur dari suatu program. Pada estimasi premi ini rangkaian langkah-langkah pembuatannya meliputi dari awal penggunaan dari program atau dengan mengawalinya dengan *start*, kemudian menuju pada *input username* dan *password* rangkaian pertanyaan untuk cek status apakah benar atau salah. Jika ya maka program tersebut akan melanjutkan pengisian data nasabah, premi, dan pembayaran premi kemudian masuk kedalam proses yakni rekap data, *cleaning*, *clustering*, hasil analisa dan prediksi yang akhirnya akan memberitahu estimasi premi. tetapi jika pilihannya tidak maka program tidak dapat berjalan.

Penggambaran *flowchart* dari estimasi premi oleh admin dapat dilihat pada gambar 3.



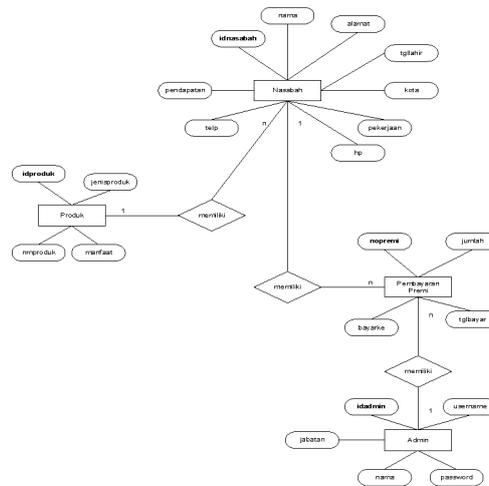
Gambar 3 Diagram Alir Sistem (Flowchart System)

Pada Gambar 3 diatas merupakan Diagram Alir Sistem (DAD) yang dimulai dari proses *login* input *username*, dan *password* jika benar maka akan admin akan melakukan pengolahan data nasabah yang kemudian nantinya akan di proses dalam bentuk *clustering* dan hasil analisa yang kemudian hasil akhirnya adalah mendapatkan hasil estimasi premi untuk nasabah.

### 3.6 Entity Relationship Diagram (ERD)

Perancangan *Entity Relationship Diagram* (ERD) bertujuan untuk memudahkan pemahaman tentang hubungan (relasi) antar tiap tabel yang ada dalam basis data.

Penggambaran *Entity Relationship Diagram* dapat dilihat pada Gambar 4 memuat informasi mengenai *Entity Relationship Diagram* sistem.



Gambar 4 *Entity Relationship Diagram*

Adapun keterangan notasi-notasi nya adalah sebagai berikut:

- n adalah many
- 1 adalah one

Pada Gambar 4 diketahui bahwa nasabah mempunyai (idnasabah, nama, alamat, tglahir, kota, pekerjaan, hp, telp, pendapatan) yang dimana nasabah adalah n (banyak) mempunyai 1 produk (nmproduk, idproduk, manfaat, dan jenis produk) kemudian nasabah tersebut melakukan transaksi pembaruan premi kepada *admin*. Pembayaran premi yang mempunyai (nopremi, tglbayar, bayarke, dan jumlah) selanjutnya *admin* mempunyai (nama, idadmin, password, username, dan jabatan) yang dimana *admin* berfungsi sebagai petugas yang melakukan transaksi pembayaran premi nasabah.

#### 4. Implementasi Sistem

Pada bagian implementasi sistem merupakan bagian yang mendeskripsikan suatu sistem agar sistem tersebut siap untuk dioperasikan. Bagian ini dilakukan setelah setelah tahap analisis data tahap perancangan aliran data dan menu sistem selesai. Tahap implementasi ini akan membahas mengenai desain *login*, *input* data, desain proses data, desain *output* sistem serta analisis hasil yang berupa pengujian sistem tersebut secara manualnya.

##### Menu Login

Merupakan tampilan *Login* merupakan tampilan awal saat sistem pertama kali dijalankan. Tampilan dapat dilihat pada Pada Gambar 5



Gambar 5 Tampilan Halaman Login

### Implementasi Estimasi Premi

Tampilan Halaman Estimasi premi merupakan halaman yang dirancang untuk menentukan premi yang akan dibayarkan oleh nasabah kepada pihak asuransi. Tampilan gambar dapat dilihat Pada Gambar 6



Gambar 6 Tampilan Hasil Analisa

### Analisa Hasil

Setelah data pendapatan nasabah dikelompokkan, kemudian dilakukan proses interpolasi. Tabel 4 merupakan tabel koefisien untuk pencarian jumlah data pendapatan nasabah.

Tabel 4 Tabel Koefisien

i	x	y	ST 1	ST 2
0	x <sub>1</sub>	y <sub>1</sub>	$(y_1 - y_0) / (x_1 - x_0) = A$	$(B - A) / (x_2 - x_0) = D$
1	x <sub>2</sub>	y <sub>2</sub>	$(y_2 - y_1) / (x_2 - x_1) = B$	
2	x <sub>3</sub>	y <sub>3</sub>	$(y_3 - y_2) / (x_3 - x_2) = C$	

Adapun keterangan notasi-notasi nya adalah sebagai berikut:

Keterangan :

x adalah pendapatan nasabah

y adalah premi nasabah

b<sub>0</sub> adalah nilai premi

i adalah jumlah data

ST adalah koefisien

Contoh :

Setelah selesai mengelompokkan data pendapatan, maka selanjutnya diketahui masing-masing pendapatan nasabah adalah  $x_1$  adalah Rp.1.500.000,  $x_2$  adalah Rp.1.750.000 dan Rp.2.000.000. Dengan premi untuk  $y_1$  adalah Rp.100.000,  $y_2$  adalah Rp.200.000, dan  $y_3$  adalah Rp.350.000. Hitunglah hasil estimasi premi nasabah tersebut ?

Penyelesaian Kasus :

Tabel 3.1 Tabel penghitungan jumlah data

i	Pendapatan Nasabah	Premi Nasabah
0	Rp 1.500.000	Rp 100.000
1	Rp 1.750.000	Rp 200.000
2	Rp 2.000.000	Rp 350.000

Rumus yang digunakan

$$i = i - 1 \dots \dots \dots (1)$$

Maka  $i = 3 - 1 = 2$

Selanjutnya menghitung ST1

**Hitung ST1 :**

$$ST = (y_1 - y_0) / (x_1 - x_0) = A \dots \dots \dots (2)$$

$$0.4 = (Rp\ 200.000 - Rp\ 100.000) / (Rp\ 1.750.000 - Rp\ 1.500.000)$$

$$ST = (y_2 - y_1) / (x_2 - x_1) = B \dots \dots \dots (3)$$

$$0 = (Rp\ 200.000 - Rp\ 200.000) / (Rp\ 2.000.000 - Rp\ 1.750.000)$$

**Hitung ST2 :**

Untuk menghitung ST2

$$ST = (B - A) / (x_2 - x_0) = C \dots \dots \dots (4)$$

$$-8.10^{-7} = (0 - 0.4) / (Rp\ 2.000.000 - Rp\ 1.500.000).$$

**Menghitung Nilai Interpolasi**

Untuk menaksir (mengestimasi) nilai di antara titik- titik data yang tepat Bentuk umum persamaan polinomial order  $n$  adalah sebagai berikut,

$$f_n(x) = b_0 + b_1(x - x_0) + \dots + b_n(x - x_0)(x - x_1) \dots (x - x_{n-1})$$

Seperti yang dilakukan interpolasi linier dan kuadrat, titik-titik data dapat dilakukan dengan evaluasi koefisien  $b_0, b_1, \dots, b_n$ . Untuk polinomial order  $n$ , diperlukan  $(n + 1)$  titik data  $x_0, x_1, x_2, \dots, x_n$ . Dengan menggunakan titik-titik data tersebut, maka persamaan berikut digunakan untuk mengevaluasi koefisien  $b_0, b_1, \dots, b_n$ .

**Nilai Koef untuk masing-masing  $b_0, b_1, \dots$  :**

Nilai  $B_0 = Rp. 100.000$  , Nilai  $B_1 = 0.4$  , Nilai  $B_2 = -8.10^{-7}$ ,

### Menghitung nilai $f(x)$ dimana $x = \text{Rp } 1.500.000$

$$f_n(x) = b_0 + b_1(x - x_0) + b_n(x - x_0)(x - x_1)$$

$$f(x) = \text{Rp } 100.000 + 0.4(\text{Rp } 1.500.000 - \text{Rp } 1.500.000) + (-8 \cdot 10^{-7})(\text{Rp } 1.500.000 - \text{Rp } 1.500.000)(\text{Rp } 1.500.000 - \text{Rp } 1.750.000)$$

$$f(x) = \text{Rp } 100.000$$

Dari hasil perhitungan didapatkan bahwa untuk pendapatan sebesar Rp.1.500.000 estimasi premi untuk nasabah diperoleh Rp.100.000/bulan.

## 5. Kesimpulan

1. Aplikasi estimasi premi berdasarkan pendapatan nasabah asuransi prudential Yogyakarta yang dibuat menggunakan metode *clustering* dan *interpolasi*.
2. Dengan adanya tampilan antar muka pengguna/*user interface* maka *user* menjadi lebih mudah dalam mengoperasikan sistem yang telah dibuat dalam tugas akhir ini.
3. Estimasi nilai premi menggunakan proses *clustering* dan *interpolasi* telah sesuai yang diharapkan untuk membantu *user* (pegawai prudential).

## Saran

Beberapa saran yang dapat diberikan untuk pengembangan aplikasi ini di masa yang akan datang, antara lain:

Sebaiknya program aplikasi ini dikembangkan dengan situs *online* jadi memudahkan nasabah untuk mengetahui berapa estimasi premi yang dibayarkan nasabah sesuai dengan pendapatan nasabah

## Referensi

- [1] Agus Setiawan, Pengantar Numerik, Andi Offset, Yogyakarta, 2010.
- [2] Al-Bahra, Ladjamudin, Analisis dan Desain Sistem Informasi, Graha Ilmu, Yogyakarta, 2005.
- [3] Furrar Utdirartatmo, *Clustering*, Andi Offset, Yogyakarta, 2009.
- [4] Hall, Brown, *Core Web Programming*, The Sun Microsystem Press, USA, 2001
- [5] John H. Mathews, *Numerical Methods for Mathematics, Science, And Engineering (second edition)*, Prentice Hall International Editions, USA, 2004.
- [6] Nugroho, Bunafit, Aplikasi Pemrograman Web Dinamis dengan PHP dan MySql, Gava Media, Yogyakarta, 2004
- [7] Sani, Susanto, Dedy Suryadi, Pengantar *Data Mining*, Andi Offset, Yogyakarta, 2010