

PENERAPAN ALGORITMA NAIVE BAYES UNTUK PEMILIHAN KELUARGA YANG MEMBUTUHKAN BANTUAN DALAM PROGRAM KELUARGA HARAPAN (PKH) (STUDI KASUS DI DESA KARANG ASIH, CIKARANG UTARA)

Muhtajuddin Danny, Saana Atmaja

Program Studi Teknik Informatika Sekolah Tinggi Teknologi Pelita Bangsa

utat@pelitabangsa.ac.id

Disetujui, 26 September 2018

Abstraksi

Program keluarga harapan adalah sebuah program pemerintah yg bertujuan untuk meringankan beban keluarga miskin atau hampir miskin dalam hal pangan, dalam implementasinya penerima bantuan ini masih belum optimal dikarenakan masih banyak penerima program keluarga harapan yang belum tepat sasaran. Desa Karang Asih yang terletak di Cikarang Utara, menjadi obyek penelitian ini, dengan penduduk Desa Karang Asih yang terletak di Cikarang Utara dapat dijadikan model, dengan penduduk sebanyak 60.543 jiwa, dimana hampir 9% atau 5.448 jiwa pada kondisi dibawah harapan hidup atau RTSM (Rumah Tangga Sangat Tidak Mampu) yang membutuhkan ulurantangan pemerintah agar anak-anak usia sekolah tidak putus sekolah. Data mining dapat dipakai untuk mempermudah mengatasi masalah yang belum optimal maka metode klasifikasi mampu menemukan model yang membedakan konsep atau kelas data dengan tujuan untuk dapat memperkirakan kelas-kelas dari suatu objek yang labelnya tidak diketahui oleh sebab itu, Algoritma Naive Bayes dapat memprediksi peluang dimasa depan berdasarkan pengalaman dimasa sebelumnya. Penelitian ini mengambil data sebanyak 70 data dan sebuah data uji dengan menggunakan 6 kriteria yaitu : Status PKH, Jumlah tanggungan, Kepala rumah tangga, Kondisi rumah, Jumlah penghasilan, dan Status pemilik rumah. Dari hasil pengujian sebanyak 70 sampel menunjukkan 3,5% tidak layak menerima program keluarga harapan

Kata kunci: Program Keluarga Harapan, Algoritma Naive Bayes

Abstract

The Hope Family Program is a government program that aims to ease the burden on poor or near-poor families in terms of food. In its implementation, the recipients of this assistance are still not optimal because there are still many recipients of the Hope Family program who have not been on target. Karang Asih Village, located in North Cikarang, is the object of this research, with residents of Karang Asih Village located in North Cikarang can be used as a model, with a population of 60,543 people, of which nearly 9% or 5,448 people are in conditions below life expectancy or RTSM (Rumah Very Poor Ladder) which requires a government test so that school-age children do not drop out of school. Data mining can be used to make it easier to solve problems that are not optimal so the classification method is able to find models that differentiate concepts or data classes with the aim of being able to estimate the classes of objects whose labels are unknown. Therefore, the Naive Bayes Algorithm can predict future opportunities based on previous experience. This study took 70 data and a test data using 6 criteria, namely: PKH status, number of dependents, head of household, house condition, total income, and home owner status. From the test results, 70 samples showed 3.5% did not deserve to receive the family hope program.

Keywords: Hope family program (PKH), Algoritma Naive Bayes

1. Pendahuluan

Program keluarga harapan (PKH) merupakan suatu program penanggulangan kemiskinan. Kedudukan PKH merupakan bagian dari program-program penanggulangan kemiskinan lainnya. Program keluarga harapan adalah suatu program yang memberikan bantuan tunai kepada rumah tangga sangat miskin (RSTM), jika mereka memenuhi persyaratan yang terkait dengan upaya peningkatan kualitas sumber daya manusia (SDM), yaitu pendidikan dan kesehatan.

Dalam perkembangan teknologi saat ini penerapan suatu sistem teknologi informasi untuk membantu dan mempermudah dalam kehidupan manusia saat ini begitu besar, teknologi informasi juga dapat diharapkan untuk proses dalam pengambilan keputusan dari suatu masalah, baik masalah yang sederhana maupun masalah yang kompleks. Pengambilan keputusan yaitu tindakan yang memilih strategi yang di yakini akan memberikan solusi terbaik.

Pelaksanaan keluarga harapan yang sudah lama terlaksanakan sejak di sahkan kebijakan bantuan tersebut oleh pemerintah. Namun berdasarkan data terakhir, RTSM peserta PKH masih merupakan

persentase yang kecil dari jumlah RTSM yang tersebar, artinya negara belum mampu menjalankan fungsi perlindungan sosial secara optimal. Pendataan RTSM masih menggunakan sistem manual. sementara pada saat yang bersamaan yang justru tidak masuk atau tidak terdata sebagai calon peserta PKH.

Didalam melakukan sebuah seleksi program keluarga harapan pastinya akan mengalami kesulitan apabila tidak menggunakan suatu metode tertentu, karena banyaknya masyarakat yang membutuhkan program pemerintah tersebut. Adanya beberapa kriteria yang digunakan untuk menentukan siapa yang berhak menerimanya sesuai harapan, sebagai calon penerima program keluarga harapan yang akan di terima, hanya masyarakat yang memenuhi kriteria – kriteria saja yang dapat menerima bantuan tersebut, untuk itu perlukan suatu metode yang dapat membantu keakuratan dalam memilih calon penerima program keluarga harapan

Desa Karang Asih yang terletak di Cikarang Utara dapat dijadikan model, dengan penduduk sebanyak 60.543 jiwa, dimana hampir 9% atau 5.448 jiwa pada kondisi dibawah harapan hidup, menurut BPS dengan penghasilan pada criteria RTSM (Rumah Tangga Sangat Tidak Mampu). Diharapkan dengan program ini dapat mengurangi putus sekolah.

2. Tinjauan Studi

2.1. Data Mining

Menurut Chairul Fadlan, Selfia Ningsih, Agus Perdana Windarto (2015) *data mining* merupakan proses ataupun kegiatan untuk mengumpulkan data yang berukuran besar kemudian mengekstraksi data tersebut menjadi informasi-informasi yang nantinya dapat digunakan. Sedangkan menurut Aziz, Hadi, & Moechammad, (2015) *data mining* adalah suatu proses yang menggunakan teknik statistik, matematika, kecerdasan tiruan, dan *machine-learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai *database* besar. Istilah *data mining* memiliki hakikat sebagai disiplin ilmu yang tujuan utamanya adalah untuk menemukan, menggali, atau menambang pengetahuan dari data atau informasi yang kita miliki. Selain itu *Data mining* merupakan analisis dari peninjauan kumpulan data untuk menemukan hubungan yang tidak diduga dan meringkas data dengan cara yang berbeda dengan sebelumnya, yang dapat dipahami dan bermanfaat bagi pemilik data. Ada beberapa teknik yang dimiliki *data mining* berdasarkan tugas yang bisa dilakukan, yaitu: deskripsi, estimasi, prediksi, klasifikasi, *clustering* dan asosiasi. Penelitian ini akan menggunakan fungsi klasifikasi dikarenakan sistem mengelompokkan data didasarkan pada karakteristik alternatif. Dalam klasifikasi variabel, tujuan bersifat kategorik. Istilah *data mining* dan *knowledge discovery in databases* (KDD) sering kali digunakan secara bergantian untuk menjelaskan proses penggalian informasi tersembunyi dalam suatu basis data yang besar. KDD dapat dibagi menjadi beberapa tahap proses, yaitu : pembersihan data (*data cleaning*), integrasi data (*data integration*), seleksi data (*data selection*), transformasi data (*data transformation*), proses *data mining*, evaluasi pola (*pattern evaluation*), dan presentasi pengetahuan (*knowledge presentation*).

2.2. Metode Algoritma Naive Bayes

Naive Bayes merupakan teknik prediksi berbasis probabilitas sederhana yang berdasarkan pada penerapan teorema bayes dengan asumsi independensi yang kuat. Dengan kata lain, dalam Naive Bayes menggunakan model fitur independen, maksud independen yang kuat pada fitur adalah bahwa data tidak berkaitan dengan data yang lain dalam kasus yang sama ataupun atribut yang lain.

Persamaan dari Teorema Bayes adalah :

$$P(H|X) = \frac{P(X|H).P(H)}{P(X)}$$

Keterangan :

X : Data sampel dengan kelas (label) yang tidak diketahui.

H : Hipotesa bahwa X adalah data dengan class (label).

P (H | X) : Probabilitas H berdasarkan kondisi X. P (H) = Peluang dari hipotesa H.

P(X|H). P (X) : Peluang data sampel X, bila diasumsikan bahwa hipotesa benar
Peluang data sampel yang diamati.

2.3. Unified modelling language (UML)

Unified Modeling Language (UML) adalah sebuah bahasa yang berdasarkan grafik atau gambar untuk memvisualisasikan menspesifikasikan, membangun dan pendokumentasian dari sebuah sistem pengembangan perangkat lunak berbasis Objek (*Object Oriented programming*). *Object-oriented programming* merupakan paradigma pemrograman yang berorientasikan kepada objek. Semua data dan fungsi didalam paradigma ini dibungkus dalam kelas - kelas atau objek - objek. Bandingkan dengan logika pemrograman terstruktur. Setiap objek dapat menerima pesan, memproses data, dan mengirim pesan ke objek lainnya.









Model data berorientasi objek dikatakan dapat memberifleksibilitas yang lebih, kemudahan mengubah program, dan digunakan luas dalam teknik piranti lunak skala besar. Lebih jauh lagi, pendukung OOP mengklaim bahwa OOP lebih mudah dipelajari bagi pemula dibanding dengan pendekatan sebelumnya dan pendekatan OOP lebih mudah dikembangkan dan dirawat.

Pada pembahasan ini akan dijelaskan 8 diagram yaitu *Use case Diagram*, *Class Diagram*, *Statechart Diagram*, *ActivityDiagram*, *Sequence Diagram*, *Collaboration Diagram*, *Component Diagram*, dan *DeploymentDiagram*. Dimana penjelasan masing-masing diagram tersebut adalah sebagai berikut:

2.3.1. Usecase Diagram

Usecase adalah salah satu diagram yang ada dalam UML (*Unified Modeling Language*). *Usecase* atau diagram *Usecase* merupakan pemodelan untuk kelakuan (*behavior*) aplikasi perangkat lunak yang akan dibuat. *Usecase* mendeskripsikan sebuah interaksi antara satu lebih actor dengan aplikasi yang akan dibuat. Secara kasar, *Usecase* digunakan untuk mengetahui fungsi/proses apa saja yang ada di dalam sebuah aplikasi dan siapa saja yang berhak menggunakan fungsi-fungsi/proses-proses itu. Syarat penamaan pada *Usecase* adalah nama didefinisikan sesimpel mungkin dan dapat dipahami.

Tabel 1. Use Case Diagram








Simbol	Nama	Keterangan
	Aktor	Menspesifikasikan himpunan peran yang pengguna mainkan ketika berinteraksi dengan use case
	Dependency	Hubungan dimana perubahan yang terjadi pada suatu elemen mandiri (independent) akan memengaruhi elemen yang bergantung padanya elemen yang tidak mandiri
	Generalization	Hubungan dimana objek anak (descendent) berbagi perilaku dan struktur data dari objek yang ada di atasnya objek induk(ancestor)
	Include	Menspesifikasikan bahwa use case sumber secara eksplisit
	Extend	Menspesifikasikan bahwa use case target memperluas perilaku dari use case sumber pada suatu titik yang diberikan
	Association	Apa yang menghubungkan antara objek satu dengan objek lainnya
System	System	Menspesifikasikan paket yang menampilkan sistem secara terbatas
	Use Case	Deskripsi dari urutan aksi-aksi yang ditampilkan sistem yang menghasilkan suatu hasil yang terukur dari suatu aktor
	Collaboration	Interaksi aturan dan elemen lain yang bekerja sama untuk menyediakan perilaku yang lebih besar dari jumlah dan elemennya (sinergi)

2.3.2. Class Diagram

Diagram kelas menggambarkan struktur aplikasi berorientasi objek dari segi pendefinisian kelas-kelas yang akan dibuat untuk membangun aplikasi. Kelas memiliki apa yang disebut atribut dan metode atau operasi. Atribut merupakan variabel-variabel yang dimiliki oleh suatu kelas, operasi atau metode adalah fungsi-fungsi yang dimiliki oleh suatu kelas.

Berikut adalah simbol-simbol yang ada di diagram kelas:

Tabel 2. Diagram Kelas

No	Gambar	Nama	Keterangan
1		Generalization	Hubungan dimana obyek anak (descendant) berbagi perilaku struktur data dari obyek yg ada diatas obyek induk (onceter)
2		Nary Association	Upaya untuk menghindari asosiasi dengan lebih dari 2 obyek
3		Class	Himpunan dari obyek2 yang berbagi atribut serta operasi yang sama
4		Collaboration	Deskripsi dari urutan aksi yang ditampilkan sistem yang menghasilkan suatu hasil terukur bagi suatu aktor
5		Relization	Operasi yang benar-benar dilakukan oleh suatu obyek
6		Depedency	Hubungan dimana perubahan yang terjadi pada satu elemen mandiri/ independent akan mempengaruhi elemen yg bergantung pada elemen tidak mandiri
7		Association	Apa yang menghubungkan antar obyek satu dengan obyek lainnya

2.3.3. Activity Diagram

Diagram aktivitas atau *activity diagram* menggambarkan *workflow* (aliran kerja) atau aktivitas dari sebuah sistem atau proses bisnis. Perlu diperhatikan di sini adalah bahwa diagram aktivitas menggambarkan aktivitas sistem bukan apa yang dilakukan aktor, jadi aktivitas yang dapat dilakukan oleh sistem.




Diagram aktivitas juga banyak digunakan untuk mendefinisikan hal- hal berikut:

1. Rancangan proses bisnis dimana setiap urutan aktivitas yang digambarkan merupakan proses bisnis sistem yang didefinisikan.
2. Urutan atau pengelompokkan tampilan dari sistem/*user interface* dimana setiap aktivitas dianggap memiliki sebuah rancangan antar muka tampilan.
3. Rancangan pengujian dimana setiap aktivitas dianggap memerlukan sebuah pengujian yang perlu didefinisikan kasus ujinya.

Berikut adalah simbol-simbol yang ada pada diagram aktivitas

Tabel 3. Activity Diagram

Gambar	Nama	Fungsi
	State	Menambahkan state untuk suatu obyek
	Activity	Menambahkan aktifitas baru pada diagram
	Start State	Memperlihatkan dimana aliran kerja mulai
	End State	Memperlihatkan dimana aliran kerja berakhir

	State Transition	Menambahkan transisi dari satu aktifitas ke aktifitas yang lain
	Transition to Self	Menambahkan transisi rekursif
	Horizontal Synchronization	Menambahkan sinkronisasi horizontal pada diagram
	Vertical Synchronization	Menambahkan sinkronisasi vertikal pada diagram
	Decision Points	Menambahkan titik keputusan pada aliran kerja

2.3.4. Sequence Diagram

Diagram sekuen menggambarkan kelakuan objek pada *usecase* dengan mendeskripsikan waktu hidup objek dan *message* yang dikirimkan dan diterima antar objek. Oleh karena itu untuk menggambar diagram sekuen harus diketahui objek-objek yang terlihat dalam sebuah *usecase* beserta metode-metode yang dimiliki kelas yang diinstansikan menjadi objek itu.

Banyaknya diagram sekuen yang harus digambar adalah sebanyak pendefinisian *usecase* yang memiliki proses sendiri atau yang penting semua *usecase* yang telah didefinisikan interaksi jalannya pesan sudah dicakup pada diagram sekuen sehingga semakin banyak *usecase* yang didefinisikan maka diagram sekuen yang harus dibuat jugasemakin banyak.

Berikut adalah simbol-simbol yang ada pada diagram sequence:

Tabel 4. Sequence Diagram

Gambar	Nama	Fungsi
	Object	Menambahkan objek baru pada diagram
	Activity message	Menggambar pesan antar dua obyek
	Message to self	Menggambar pesan kepada diri sendiri
	Return Message	Menggambar pengembalian dari pemanggilan prosedur
	Destruction marker	Menggambarkan saat obyek tertentu di hancurkan menuju dirinya sendiri
	Return Message	Menggambarkan saat objek tertenti di hancurkan
	Destruction Marker	Menambahkan sinkronisasi horizontal pada diagram

2.4. Microsoft Visual Studio 2010

Visual basic 2010 menurut Mustakim & Krisnawati, (2013) merupakan salah satu bagian dari produk pemrograman yang dikeluarkan oleh microsoft, yaitu Microsoft Visual Studio 2010. Sebagai produk pengembangan atau *Integrated Develotmen Environment* (IDE) andalan yang dikeluarkan oleh microsoft, visual studio 2010 berisi beberapa IDE pemrograman seperti Visual Basic, Visual C++, Visual Web Develover, Visual C#, Dan Visual F#. Semua IDE tersebut sudah mendukung penuh implementasi .Net Framework terbaru yaitu .Net Fram Work 4.0 yang mrupakan pengembangan dari . Net Framework 3.5.

2.5. Xampp

Xampp berdasarkan penelitian Randi, Yulie , & Arie, (2015) adalah perangkat lunak bebas, mendukung banyak sistem operasi, merupakan kompilasi dari beberapa program. Fungsinya adalah sebagai server yang berdiri sendiri (localhost) yang terdiri atas program apache HTTP server, MYSQL database, dan penerjemah bahasa yang ditulis dengan bahasa pemrograman PHP dan perl. Nama xampp yaitu singkatan dari X (empat sistem operasi apapun). Apache, mysql, php dan perl. . program ini tersedia dalam GNU *general public licance* dan bebas, merupakan web server yang mudah digunakan yang dapat melayani tampilan halaman web yang dinamis. Untuk mendapatkannya dapat mendownload langsung dari web resminya.

2.6. Mysql

MYSQL adalah sebuah implementasi dari sistem manajemen basisdata relasional (RDBMS) berdasarkan penelitian Henry & Eri, (2012) yang didistribusikan secara gratis dibawah lisensi GPL (*General Public License*). Setiap pengguna dapat secara bebas menggunakannya, namun dengan batasan perangkat lunak tersebut tidak dapat dijadikan produk turunan yang bersifat komersial. MYSQL sebenarnya merupakan turunan salah satu konsep utama dalam basisdata yang telah ada sebelumnya ; SQL (*Structured Query Language*). SQL adalah sebuah konsep pengoprasian basisdata, terutama untuk pemilihan satu seleksi dan pemasukan data, yang memungkinkan dalam pengoprasian data dikarjakan dengan mudah secara otomatis.

Kendala sistem basisdata DBMS dapat diketahui dari cara kerja pengoptimasinya dalam melakukan proses perintah-perintah SQL yang dibuat pengguna maupun program-program aplikasi yang memanfaatkannya. Sebagai peladan basisdata, MYSQL mendukung operasi basisdata transaksional maupun basisdata non transaksional, pada modus operasi non transaksional, MYSQL dapat dikatakan unggul dalam hal unjuk kerja dibandingkan perangkat peladen basisdata kompetitor lainnya.

2.7. Cristal Report

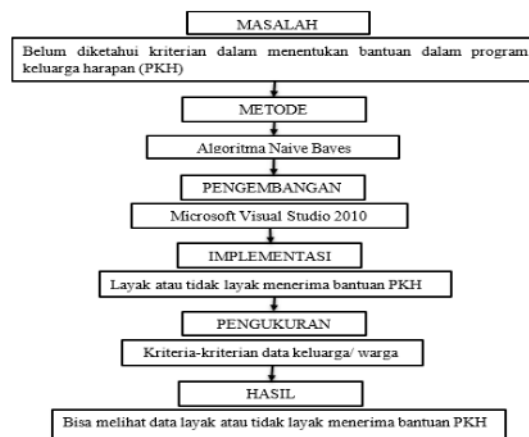
Cristal report menurut Triana & Stephanie, (2015) merupakan peranti standard untuk pembuatan laporan pada sistem operasi windows, dimana cetakan/template laporan yang dihasilkan dapat disertakan pada banyak bahasa pemrograman.

Crystal report terdiri dari tiga bagian utama, yaitu :

1. *Toolbox*, yang berfungsi untuk menambahkan objek-objek kedalam repot designer.
2. *Field exploler*, yang befungsi untuk menampilkan daftar field, formula, dan pernyataan-pernyataan sql serta yang lainnya.
3. *Report designer*, yang berfungsi untuk meletakkan objek-objek yang digunakan pada laporan.

3. Kerangka Konsep

Kerangka konsep dalam penelitian adalah dimulai dari penetapan masalah, objek masalah, kriteria, kemudian tindak lanjut berikutnya adalah perancangan sistem menggunakan diagram use case, activity diagram, sequence diagram setelah itu dilanjutkan dengan membuat program pada excel untuk data training, melakukan perhitungan-perhitungan probablitas kelayakan. Seperti pada gambar 1 dihalaman berikut :



Gambar 1. Kerangka Konsep

4. Desain Penelitian/ Metodologi

4.1. Pengumpulan data

Data diambil dari data primer yang menggunakan teknik wawancara terhadap 70 orang Desa Karang Asih yang memiliki kriteria masuk dalam program keluarga harapan (pkh). Penerapan Algoritma Naive Bayes

4.2. Penggunaan Teknik Naive bayes

Teknik Naive merupakan teknik prediksi berbasis probabilitas sederhana yang berdasarkan pada penerapan teorema bayes dengan asumsi independensi yang kuat. Persamaan dari teorema bayes adalah :

$$P(H|X) = \frac{P(X|H).P(H)}{P(X)} \quad (1)$$

Dimana :

- X : Data dengan class yang belum diketahui
- H : Hipotesis data merupakan suatu class spesifik
- P(H|X) : Probabilitas hipotesis H berdasar kondisi X (posteriori probabilitas)
- P(H) : Probabilitas hipotesis H (prior probabilitas)
- P(X) : Probabilitas X

Penjabaran lebih lanjut rumus bayes tersebut dilakukan dengan menjabarkan (C|X1...Xn). menggunakan aturan perkalian sebagai berikut.

$$\begin{aligned} P(C|X_1, \dots, X_n) &= P(C) P(X_1, \dots, X_n|C) \\ &= P(C)P(X_1|C)P(X_2, \dots, X_n|C, X_1) \\ &= P(C)P(X_1|C)P(X_2|C, X_1)P(X_3, \dots, X_n|C, X_1, X_2) \\ &= P(C)P(X_1|C)P(X_2|C, X_1)P(X_3|C, X_1, X_2)P(X_4, \dots, X_n|C, X_1, X_2, X_3) \\ &= P(C)P(X_1|C)P(X_2|C, X_1)P(X_3|C, X_1, X_2) \dots P(X_n|C, X_1, X_2, X_3, \dots, X_{n-1}) \dots (2) \end{aligned}$$

Akibat perhitungan semakin sulit untuk dilakukan dan disinilah digunakan asumsi independensi seperti dibawah ini:

$$\begin{aligned} P(X_i|X_j) &= \frac{P(X_i \cap X_j)}{P(X_j)} \\ &= \frac{P(X_i)P(X_j)}{P(X_j)} \\ &= P(X_i) \end{aligned}$$

Untuk $i \neq j$, sehingga

$$P(X_i|C, X_j) = P(X_i|C) \quad \dots(3)$$

Selanjutnya Penjabaran (P(C|X1.....Xn) dan dapat di sederhanakan menjadipersamaan :

$$\begin{aligned} P(C|X_1, \dots, X_n) &= \frac{P(X_2|C)P(X_3|C) \dots}{P(X_1|C)} \\ &= \prod_{i=1}^n P(X_i|C) \quad \dots(4) \end{aligned}$$

Keterangan :

$$\prod_{i=1}^n P(X_i|C) = \text{perkalian ranting antar atribut}$$

Untuk klasifikasi dengan data continue atau data angka menggunakan rumus distribusi Gaussian dengan 2 parameter : mean dan varian :

$$P(X_i=x_i|C=c_j) = \frac{1}{\sqrt{2\pi\sigma_{ij}}} \exp\left(-\frac{(x_i-\mu_{ij})^2}{2\sigma_{ij}^2}\right) \quad 5$$

- Dimana :
- P : Peluang
 - X_i : Atribut ke i
 - X_j : Nilai atribut ke i
 - C : Kelas yang dicari
 - C_i : Sub kelas Y yang dicari
 - μ : menyatakan rata-rata seluruh atribut
 - σ : Deviasi standar,

Dalam metode Naive Bayes diperlukan data latih dan data uji yang ingin diklasifikasikan dalam Naive Bayes, semakin baik hasil yang prediksi yang diberikan dalam menghitung menghitung P(C_i)

$$P(c_i) = \frac{s_i}{s} \dots(6)$$

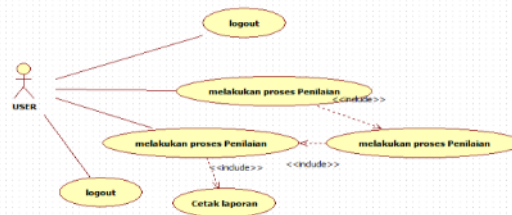
Dimana s_i adalah jumlah data training dari kategori C_i dan s P(X_i|C_i) yang merupakan probabilitas posterior X

4.3. Perancangan Sistem

Pada tahapan ini akan dilakukan perancangan terhadap sistem yang diusulkan. Adapun perancangan dari sistem ini dapat di gambarkan dengan Unifiel Language Modeling (UML). Yang meliputi usecase diagram, class diagram, actifitydiagram, dan sequence diagram.

4.3.1. Usecase Diagram

Use case diagram digunakan untuk memodelkan fungsionalitas- fungsionalitas sistem atau perangkat lunak dilihat dari pengguna yang ada diluar sistem. Use case pada dasarnya merupakan unit fungsionalitas koheren yang diekspresikan sebagai transaksi-transaksi yang terjadi antara aktor dan sistem berikut ini.



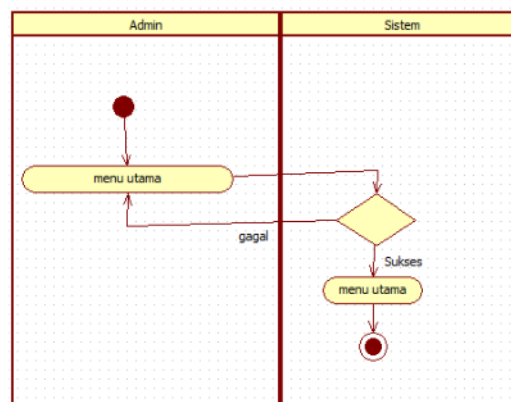
Gambar 2. Sistem pemilihan penerimaan calon PKH

4.3.2. Activity Diagram

Berikut adalah beberapa diagram aktivitas yang terbentuk dari kegiatan yang pada sistem yang diusulkan:

4.3.2.1. Login

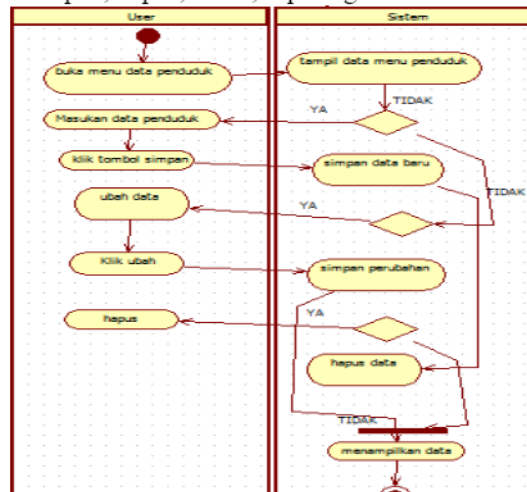
Login ini yaitu gambaran aktivitas aktor untuk masuk kedalam sistem. Setelah aktor memasukan username dan password maka sistem akan mengevaluasi apakah username dan password sudah benar atau belum, jika salah maka sistem akan kembali ke halaman login untuk memasukan kembali password dan username, Jika benarsistem akan menampilkan halaman utama.



Gambar 2. Activity diagram login

4.3.2.2. Activity Diagram Data Warga

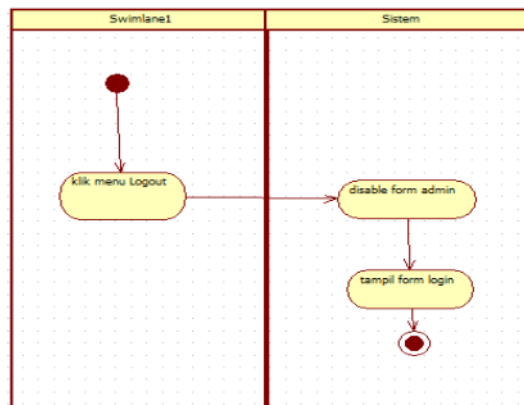
Diagram Activity ini berfungsi untuk menjelaskan cara melakukan pengolahan data warga/penduduk sesuai dengan kebutuhan, seperti sipan, hapus, ubah, seperti gambar berikut.



Gambar 3. Activity diagram data Penduduk

4.3.2.3. Activity Diagram Data Warga (log-out)

Activitas ini bertujuan aktor keluar dari sistem yang sedang berjalan .



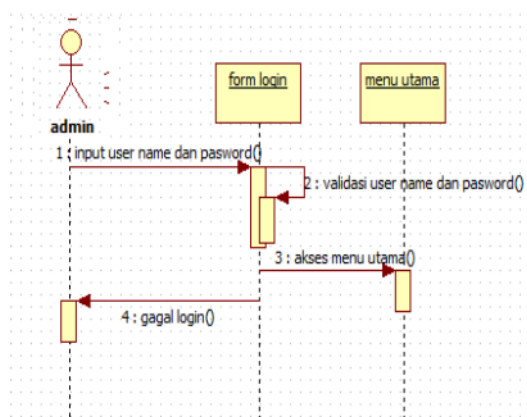
Gambar 4. Activity diagram logout

4.3.2.4. Sequence diagram

Berikut ini adalah beberapa diagram sequence yang terbentuk dari kegiatan yang pada sistem yang diusulkan :

1. Sequence Diagram Login

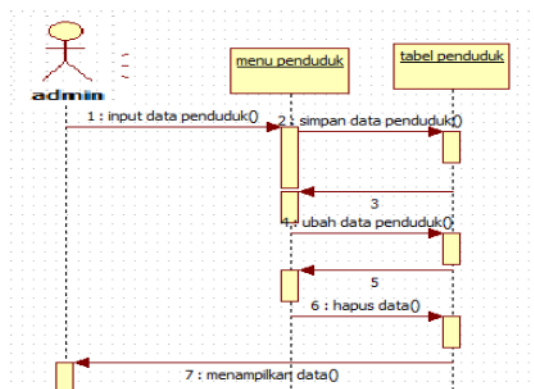
Sequence Login yaitu sebuah gambaran untuk masuk kedalam sistem. Setelah aktor memasukan username dan pasword maka sistem akan memverifikasi apakah user name dan password sudah benar, jika benarmaka sistem akan menampilkan menu utama.



Gambar 5. Sequence Diagram Login

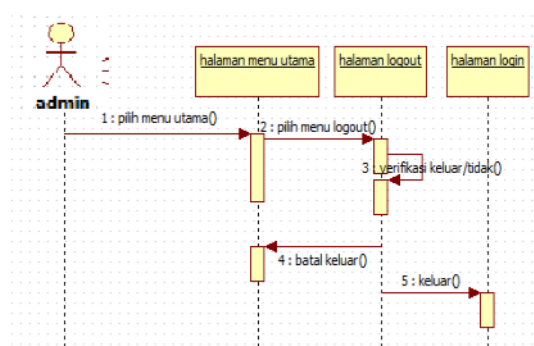
2. Sequence Data Penduduk

Pada diagram ini menggambarkan interaksi admin dengan sistem dan database dalam melakukan pengolahan data pelajar.



Gambar 6. Sequence Data Penduduk

3. Logout
Aktivitas ini bertujuan aktor keluar dari sistem



Gambar 7. Sequence Logout

5. Hasil Penelitian dan Pengujian

5.1.1. Hasil Penelitian

Kegiatan ini dimulai dengan menetapkan data awal yang berguna sebagai objek penelitian, data tersebut sebagai berikut :

Tabel 2. Data Training

No	Nama Kepala Rumah Tangga	PKH	Jumlah Tanggungan	Kepala Rumah Tangga	Kondisi Rumah	Jumlah Penghasilan	Status Pemilik Rumah	Status
1	Watiem	non	1	Perempuan	Batu	rendah	Milik Sendiri	Layak
2	Tugimin	non	4	Laki-Laki	Bambu	sedang	Milik Sendiri	Layak
3	Samino	non	3	Laki-Laki	Batu	Tinggi	Milik Sendiri	Tidak Layak
4	Mesman K	non	2	Laki-Laki	Batu	sedang	Milik Sendiri	Tidak Layak
5	Arianto	1	2	Laki-Laki	Papan	rendah	Milik Sendiri	Layak
6	Asani	non	3	Laki-Laki	Batu	sedang	Milik Sendiri	Tidak Layak
7	Sumiyadi	non	3	Laki-Laki	Bambu	rendah	Milik Sendiri	Layak
8	Ponidi	1	5	Laki-Laki	Bambu	sedang	Milik Sendiri	layak
9	Jabar Damanik	1	9	Laki-Laki	Batu	sedang	Milik Sendiri	layak
10	Safruddin	non	4	Laki-Laki	Batu	Tinggi	Milik Sendiri	Tidak Layak
11	Bonim	non	1	Perempuan	Batu	rendah	Milik Sendiri	Layak
12	paisal	non	0	Laki-Laki	Batu	rendah	sewa	layak

13	Nuri	non	1	Perempuan	Batu	rendah	Milik Sendiri	Layak
14	adam Husein	non	2	Laki-Laki	Batu	Tinggi	Milik Sendiri	Tidak Layak
15	nani	non	2	Perempuan	Batu	Tinggi	Milik Sendiri	Tidak Layak
16	Suroso	non	2	Laki-Laki	Batu	sedang	Milik Sendiri	Layak
17	Majid	non	1	Laki-Laki	Papan	rendah	Milik Sendiri	Layak
18	Mulyadi	non	3	Laki-Laki	Batu	rendah	sewa	layak
19	Jafar	non	2	Laki-Laki	Batu	rendah	Milik Sendiri	Tidak Layak
20	Simon	non	5	Laki-Laki	Papan	rendah	Milik Sendiri	Layak
70	Damanik	non	2	Laki-Laki	Batu	Tinggi	Milik Sendiri	Tidak Layak
71	Fajar	non	4	Laki-Laki	Batu	sedang	Milik Sendiri	?

Tahap awal proses penghitungan Naive Bayes adalah dengan melakukan pengambilan data training dari data yang telah diperoleh. Variabel yang digunakan dala klasifikasi kelayakan masyarakat penerima program PKH yaitu :

1. Nama kepala keluarga
2. Setatus PKH yang dikategorikan menjadi 2 yaitu penerima PKH (1) dan tidak menerima PKH (2).
3. Kepala rumah tangga yang dikelompokan menjadi 2 kategori yaitu laki-laki dan perempuan dikarenakan suami atau kepala rumah tangga laki- laki telah meninggal dunia.
4. Kondisi rumah yang di kelompokan menjadi 3 kategori yaitu batupermanen, bambu anyam dan papan disesuaikan dengan keadaan kondisi rumah yang sebenarnya.
5. Jumlah penghasilan dikategorikan menjadi 3 yaitu rendah, sedang, tinggi

Tabel 3. Jumlah penghasilan

Kriteria	Kategori	Keterangan
Jumlah Penghasilan	Rendah	< Rp. 1.500.000
	Sedang	Rp.1.500.000 s/d Rp.2.500.000
	Tinggi	> Rp.2.500.000

5.1.2. Perhitungan Probabilitas Prior (P(Ci))

Dari 70 data latih yang digunakan diketahui kelas C0 (Tidak Layak) sebanyak 41 data. Perhitungan probabilitas prior kemungkinan tidak layak menerima PKH dapat dilakuka dengan persamaan yaitu :

$$P(C_0) = \frac{29}{70} = 0.41$$

Sedangkan penghitungan probabilitas kemungkinan layak yaitu :

$$P(C_0) = \frac{41}{70} = 0.59$$

5.1.3. Perhitungan probabilitas

Untuk menghitung setiap kemungkinan hasil kriteria P (XX Jumlah tanggungan Ci dapat dilihat pada tabel berikut :

Tabel 4. Probabilitas setatus PKH

Status PKH	Jumlah Kejadian		Probabilitas	
	Layak	Tdk Layak	Layak (C1)	Tdk Layak (C0)
Non	35	29	0,853585	1
1	6	0	0,146415	0
Jumlah	41	29	1	1

Untuk menghitung setiap kemungkinan hasil kriteria P(XX Jumlah Tanggungan Ci dapat dilihat

pada tabel berikut :

Tabel 5. Probabilitas jumlah tanggungan

Jumlah Tanggungan	Jumlah Kejadian "Dipilih"		Probabilitas	
	Layak	Tdk Layak	Layak (C1)	Tdk Layak
0	2	0	0.048780	0
1	7	2	0.170732	0.068966
2	12	11	0.292683	0.379310
3	11	10	0.268293	0.344828
4	4	6	0.097561	0.206897
5	2	0	0.048780	0
>5	3	0	0.073171	0
Jumlah	41	29	1	1

Untuk menghitung setiap kemungkinan hasil kriteria P (XX Kepala Rumahtangga Ci dapat dilihat pada tabel berikut :

Tabel 6. Probabilitas kepala rumah tangga

Jumlah Tanggungan	Jumlah		Probabilitas	
	Layak	Tdk Layak	Layak (C1)	Tdk Layak (C0)
Perempuan	13	1	0.317073	0.034483
Laki	28	28	0.682927	0.965517
Jumlah	41	29	1	1

Untuk menghitung stiap kemungkinan hasil kriteria P (XX Kondisirumah Ci dapat dilihat pada tabel berikut :

Tabel 7. Probabilitas kondisi rumah

Kondisi Rumah	Jumlah Kejadian		Probabilitas	
	Layak	Tdk Layak	Layak (C1)	Tdk Layak (C0)
Bambu	12	0	0.292683	0
Papan	11	3	0.268293	0.103448
Batu Permanen	18	26	0.439024	0.896552
Jumlah	41	29	1	1

Untuk menghitung setiap kemungkinan hasil kriteria P (XX Jumlah penghasilan Ci dapat dilihat pada tabel berikut :

Tabel 8. Probabilitas jumlah penghasilan

KondisiRumah	Jumlah		Probabilitas	
	Layak	Tdk Layak	Layak (C1)	Tdk Layak
Rendah	29	3	0.707317	0.103448
Sedang	11	16	0.268293	0.551724
Tinggi	1	10	0.024390	0.344828
Jumlah	41	29	1	1

Untuk menghitung setiap kemungkinan hasil kriteria P (XX Statusrumah Ci dapat dilihat pada tabel berikut :

Tabel 9. Probabilitas Status pemilik rumah

Status Pemilik Rumah	Jumlah Kejadian "Dipilih"		Probabilitas	
	Layak	Tdk Layak	Layak (C1)	Tdk Layak (C0)
Sewa	4	3	0.097561	0.103448
Milik Sendiri	37	26	0.902439	0.896552

Jumlah	41	29	1	1
--------	----	----	---	---

5.1.4. Perhitungan manual

Berikut ini perhitungan manual dengan menggunakan data uji yang dapat dilihat pada tabel berikut dengan menggunakan metode Naive Bayes.

Tabel 10. Data uji

NO	Nama kepala rumah tangga	PKH	Jumlah tanggungan	Kepala rumah tangga	Jumlah penghasilan	Status pemilik rumah	Status
71	Fajar	non	4	Laki-laki	Batu permanen	Milik sendiri	?

5.2. Pengujian

5.2.1. Pendefinisian Variabel

Berdasarkan tabel 10 dapat didefinisikan data uji X adalah sebagai berikut:

$X = \{X \text{ Status PKH} = \text{Non}, X \text{ jumlah Tanggungan} = 4, X \text{ Kepala Rumah tangga} = \text{laki-laki}, X = \text{Kondisi Rumah} = \text{Batu Permanen}, X \text{ Jumlah Penghasilan} = \text{Sedang}, X \text{ Status Pemilik Rumah} = \text{Milik Sendiri}\}$

5.2.2. Pendefinisian Probabilitas Prior P(Ci)

Untuk kelas tidak layak (C0) sebesar 0.41 dan kelas layak (CI) sebesar 0.59.

5.2.3. Perhitungan Probabilitas Data uji

Dilakukan klasifikasi kedalam kelas tidak layak (C0) Penerima PKH dengan ketentuan nilai masing-masing kriteria yaitu : 1, 0.206897, 0.965517, 0.896552, 0.551724, 0.896552, kemudian dari

masing-masing nilai tersebut dikalikan :

$$P(X|C0) = P(X \text{ Status PKH} = \text{Non}|C0) \times P(X \text{ Jumlah Tanggungan} = 4|C0) \times P(X \text{ Kondisi Rumah} = \text{Batu Permanen}|C0) \times P(X \text{ Jumlah Penghasilan} = \text{Sedang}|C0) \times P(X \text{ Status Pemilik Rumah} = \text{Milik Sendiri}|C0)$$

$$= 1 \times 0.206897 \times 0.965517 \times 0.896552 \times 0.896552$$

$$= 0.088590$$

Untuk menghitung klasifikasi kedalam kelas layak (CI) yaitu : 0.853659, 0.097561, 0.682927, 0.439024, 0.268293, 0.902439

Kemudian dari masing-masing kriteria tersebut dikalikan

$$P(X|CI) = P(X \text{ Status PKH} = \text{Non}|CI) \times P(X \text{ Jumlah Tanggungan} = 4|CI) \times P(X \text{ Kondisi Rumah} = \text{Batu Permanen}|CI) \times P(X \text{ Jumlah Penghasilan} = \text{Sedang}|CI) \times P(X \text{ Status Pemilik Rumah} = \text{Milik Sendiri}|CI)$$

$$= 0.853659 \times 0.097561 \times 0.682927 \times 0.439024 \times 0.268293 \times 0.902439$$

$$= 0.006046$$

5.2.4. Pemaksimalan P(X|Ci) P (Ci)

Perhitungan pemaksimal untuk klasifikasi kelas tidak layak (C0) adalah dengan cara mengalikan $P(X|C0)$ dengan $P(C0)$:

$$P(C0 X) = P(X|C0) \times P(C0)$$

$$= 0.088590 \times 0.41$$

$$= 0.036322$$

klasifikasi kelas layak (CI) adalah dengan cara mengalikan $P(X|CI)$ dengan $P(CI)$

$$P(CI X) = P(X|CI) \times P(CI)$$

$$= 0.006046 \times 0.59$$

$$= 0.003567$$

Dari perhitungan diatas dapat disimpulkan $P(C0 X) = 0.036322$ sedangkan nilai $P(CI X) = 0.003567$. Berdasarkan kesimpulan hasil dari nilai tersebut dapat diambil kesimpulan bahwa $P(C0 X) > P(CI X)$ maka data uji tersebut diklasifikasikan kedalam kelas "Tidak Layak" dalam menerima program PKH.

6. Kesimpulan

Berdasarkan Penelitian diatas didapat beberapa kesimpulan sebagai berikut :

1. Dari penerapan metode yang diterapkan yaitu Algoritma Naive Bayes dapat diketahui bahwa metode Naive Bayes lebih baik kinerjanya dalam penilaian penentuan program PKH dari sejumlah data 70 calon Penerimaan program PKH dan bisa menentukan layak dan tidak layak menerima bantuan tersebut. Oleh karena itu penulis menyarankan agar melakukan penilaian pemberian program PKH dengan menggunakan metode Naive Bayes.
2. Hasil perhitungan dari 70 sampel ada 3,5% yang tidak layak menerima bantuan tersebut.
3. Diharapkan rancangan aplikasi ini dapat memberi kemudahan serta memperkecil kesalahan yang mungkin

Daftar Pustaka

- Agustian, Wildan, Dra Sri Setyaningsih, & Arie Qur'ania. 2015. "Klasifikasi Buah Jeruk Menggunakan Metode Naive Bayes Berdasarkan Analisis Tekstur Dan Normalisasi Warna."
- Auliya, R., fitrah, Y., & Mayang, S. (2015) "Pelatihan Komputer Program Microsoft Excel 2013 Pada SMAN 12 BANJARMASIN. *Jurnal- ikhsan*, 9.
- Ciputra, Antonio, Ajib Susanto, dkk. 2018. "Klasifikasi Tingkat Kematangan Buah Apel Manalagi Dengan Algoritma Naive Bayes Dan Ekstraksi Fitur Citra Digital" 9 (1): 465–72.
- Hariadi, Y . , & Nurchholis, S. (2018) "Penerapan Data Mining Sebagai Model Seleksi Penerima Beasiswa Penuh (Study Kasus STIE PERBANAS SURABAYA). *LINK*, 2-8.
- Henry , Y ., & Eri, Z. (2012: 124-132). "Perancangan Bangun Sistem Perpustakaan untuk Jurnal Elektronik. *Jurnal teknologi informasi Dinamik volume 17, no 2, 128*.
- Monalisa Sulfia.(2016). " Sistem Penyeleksi Bantuan Langsung Tunai dengan metode Naive Bayes, artikel skripsi monalisa silfia 11.1.03.02.0247". Teknik Informatika, Universitas Nusantara PGRI Kediri.
- Mustakim, & Krisnawati. (2013) "Analisis Dan perancangan Sistem Informasi Akutansi Laporan Keuangan pada Pusat Perusahaan Distributor Air Mineral Mq Jernih Yogyakarta. *Jurnal ilmiah Dasi Vol No. 04, 26-30*.
- Subhan Dika Putra, Titi Purwandi, Irlandia Ginanjar (2018) "Validasi Model Klasifikasi Program Keluarga Harapan Menggunakan Naive Bayes Clasifier" .Universitas Padjadjaran.
- Triana, E., & Stephanie, D (2015). Sistem I nformasi Pemakaian Sparepart Mesin Packing Pada PT.XYZ. *Jatisi Vol 1 No 2,167*.