IMPLEMENTASI METODE NAIVE BAYES UNTUK KLASIFIKASI TARAF HIDUP MASYARAKAT SEJAHTERA PADA KOTA BATU

SKRIPSI

Untuk memenuhi sebagian persyaratan mencapai gelar Sarjana Komputer

Disusun Oleh:

Damara Budi Setyawan

NIM. 105090603111004

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

2016

PERSETUJUAN

Implementasi Metode Naive Bayes Untuk Klasifikasi Taraf Hidup Masyarakat Sejahtera Pada Kota Batu

SKRIPSI

Untuk memenuhi sebagian persyaratan mencapai gelar Sarjana Komputer

Disusun Oleh:

Damara Budi Setyawan

NIM. 105090603111004

Skripsi ini telah diuji dan dinyatakan lulus pada

10 November 2016

Telah diperiksa dan disetujui oleh:

Dosen Pembimbing I	Dosen Pembimbing II
Candra Dewi, S.Kom, M.Sc	**Rekyan Regasari M.P, S.T, M.T**
NIP. 19771114 200312 2 001	**NIK. 2011027704142001**

Mengetahui
Ketua jurusan teknik informatika

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

PERNYATAAN ORISINALITAS SKRIPSI

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 dikutip dalam naskah ini dan disebutkan dalam sumber kutipan dan daftar pustaka.

Apabila ternyata di dalam 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, 10 November 2016
Mahasiswa,

Damara Budi Setyawan
NIM. 105090603111004

KATA PENGANTAR

Puji dan syukur penulis panjatkan kehadirat Alloh SWT, yang telah melimpahkan berkat dan rahmat-Nya kepada penulis, sehinggga penulis dapat menyelesaikan skripsi yang berjudul "Implementasi Metode Naïve Bayes Untuk Klasifikasi Taraf Hidup Masyarakat Sejahtera Pada Kota Batu".

Pada penyusunan skripsi ini tidak akan terselesaikan semata-mata hasil kerja penulis sendiri, melainkan berkat bimbingan dan dorongan dari pihak-pihak yang telah membantu, baik secara langsung maupun tidak langsung. Maka dari itu penulis ingin mengucapkan banyak terima kasih yang tak terhingga serta penghargaan yang setinggi-tingginya kepada orang-orang yang telah membantu penulis secara langsung maupun tidak langsung. Kepada yang terhormat:

1. Tuhan YME dengan rahmat-Nya penulis dapat menyelesaikan skripsi dengan baik.
2. Kedua orang tua penulis yang telah mendidik dan memberikan support kepada penulis sehingga penulis bias berjalan sejauh ini.
3. Ibu Candra Dewi, S. Kom, M. Sc selaku dosen penasehat akademik dan pembimbing 1 yang senantiasa selalu sabar memberikan nasihat dan pengarahan kepada penulis ketika mengalami kendala dalam pengerjaan.
4. Ibu Rekyan Regasari Mardi Putri, ST., MT. sekalu dosen pembimbing 2 yang aktif mengingatkan mahasiswa bimbingannya untuk selalu bersemangat dan memberi panggilan secara tidak langsung kepada seluruh mahasiswa yang masuk dalam grup WAnya.
5. Dosen dosen PTIIK dan MIPA yang telah memberikan ilmunya kepada penulis, sehingga penulis dapat menyelesaikan skripsinya dengan baik.
6. Teman teman kontrakan Kendalsari yang membantu penulis dari awal kuliah hingga akhir Alex, Aris, Bayu, Danny dan Selly.
7. Teman teman Ilmu Komputer 2010

Malang, 10 November 2016

Damara Budi Setyawan
(105090603111004)

ABSTRAK

Indonesia memiliki wilayah yang sangat luas dan merupakan negara kepulauan terbesar ke-2 di dunia, yang memiliki kekayaan alam yang sangat melimpah, dengan iklim tropis yang menyebabkan Negara Indonesia memiliki waktu produktif sepanjang hari salama setahun, dalam hal ini Negara Indonesia seharusnya memiliki kelebihan dari Negara Negara eropa atau Negara dengan empat musim, karena masa produksi yang optimal hanya 3 musim, namun hal tersebut berbanding terbalik dengan kondisi sumber daya manusianya, kesejahteraan masyarakat Indonesia masih dibilang kurang dibandingkan dengan Negara Negara empat musim tersebut (National Geographi Indonesia, 2012).

Tujuan yang ingin dicapai dari penulisan skripsi ini adalah untuk mengetahui cara melakukan implementasi metode Naïve Bayes pada kasus klasifikasi Keluarga Sejahtera. Mengetahui tingkat akurasi algoritma pembelajaran Naïve Bayes untuk klasifikasi Keluarga sejahtera

Metode yang di gunakan mengacu Naïve Bayes, Naïve Bayes merupakan teknik prediksi berbasis probabilitas sederhana yang berdasarkan pada penerapan teorema Bayes (aturan Bayes) dengan asumsi independen yang kuat (naïf) dimana diasumsukan kondisi antar fitur saling bebas, atau dengan kata lain metode ini mengasumsikan bahwa keberadaan sebuah fitur tidak ada kaitannya dengan keberadaan atribut yang lain. Data yang kita dapat dari ekstraksi fitur, terkadang tidak sesuai dengan apa yang kita prediksi diatas kertas. Pada kenyataanya selalu ada noise pada data set dari proses ekstraksi fitur, dikarenakan adanya batasan batasan pada proses ekstraksi, yang menyebabkan data set yang kita terima, tidak mutlak seperti apa yang kita harapkan (Han, 2001). Apalagi saat kita mengambil banyak sampel untuk kita ekstraksi fiturnya, maka bisa jadi noise yang kita dapatkan akan semakin banyak. Metode Naïve Bayes tidak dapat memproses suatu fitur jika ditemukan padanya missing value, yaitu kekosongan nilai pada salah satu datanya. Akan tetapi karena Metode Naïve Bayes merupakan salah satu metode yang tingkat akurasinya tinggi, ketika suatu klasifikasi yang menggunakan algoritma Naïve Bayes ini menemukan missing value, maka menurut beberapa literatur, ketika menemukan missing value, kita dapat mengisi sendiri missing value tersebut dengan menggunakan nilai rata rata dari kelas tersebut. Data yang digunakan dalam skripsi ini didapatkan dari Posyandu Kecamatan Bumiaji, Kota Batu. Data yang diambil adalah data indikator Keluarga Sejahtera, dan status TKS pada data pemutakhiran keluarga tahun 2013 di Desa/Kelurahan Tulungrejo, RT 01, 02, 03 – RW 04 dan RT 01, 02, 03, 04, 05 – RW 05, Kecamatan Bumiaji, Kota Batu.

Berdasarkan penelitian dengan judul Implementasi Metode Naïve Bayes Untuk Klasifikasi Taraf Hidup Masyarakat Sejahtera Pada Kecamatan Batu diperoleh hasil bahwa hasil uji coba skenario pertama menunjukkan bahwa 100 data latih setelah dilakukan pengujian sebanyak 55 kali mendapatkan akurasi 92,73%. Hasil uji coba skenario kedua 20 data latih, dilakukan 50 pengujian,

didapatkan akurasi 72%. Hasil uji coba skenario ketiga, dengan 10 data latih, dan dilakukan 55 kali pengujian, didapatkan akurasi sebesar 60%.

Kata kunci: Klasifikasi, Naïve Bayes, Missing data, Keluarga Sejahtera, BKKBN

DAFTAR ISI

DAFTAR TABEL

DAFTAR GAMBAR

BAB I. PENDAHULUAN

1.1 Latar Belakang

Padatnya tingkat kependudukan di Indonesia menyebabkan terjadinya masalah pembangunan kesejahteraan di Indonesia. Indonesia memiliki wilayah yang sangat luas dan merupakan Negara kepulauan terbesar ke-2 di dunia, yang memiliki kekayaan alam yang sangat melimpah, dengan iklim tropis yang menyebabkan Negara Indonesia memiliki waktu produktif sepanjang hari salama setahun, dalam hal ini Negara Indonesia seharusnya memiliki kelebihan dari Negara Negara eropa atau Negara dengan empat musim, karena masa produksi yang optimal hanya 3 musim, namun hal tersebut berbanding terbalik dengan kondisi sumber daya manusianya, kesejahteraan masyarakat Indonesia masih dibilang kurang dibandingkan dengan Negara Negara empat musim tersebut (National Geographic Indonesia, 2012). Masalah ini mengakibatkan beban pemerintah dalam menurunkan tingkat kemiskinan dan meningkatkan kesejahteraan masyarakat menjadi berat. Langkah pemerintah dalam rangka mengatasi masalah pembangunan kesejahteraan masyarakat secara menyeluruh telah menjadi pekerjaan rumah jangka panjang dari pemerintah dengan mengacu UU No. 10 tahun 1992 yang mengatur tentang perkembangan kependudukan dan pembangunan keluarga sejahtera.

Salah satu upaya pemerintah untuk mewujudkan pemerataan pembangunan kependudukan di Indonesia adalah dengan melaksanaan pendataan keluarga setiap tahunnya di seluruh wilayah Indonesia. Dalam pelaksanaannya, masyarakat Indonesia diklasifikasikan menjadi beberapa kelompok berdasarkan status kesejahteraannya yaitu tahap keluarga prasejahtera, keluarga sejahtera I, keluarga sejahtera II, keluarga sejahtera III, dan keluarga sejahtera III plus, maka dapat memberikan informasi mengenai laporan hasil pengklasifikasian atau pengelompokan penduduk, sehingga pemerintah dapat meningkatkan kualitas pelaksanaan program pembangunan keluarga sejahtera. Selain itu, hasil pengklasifikasian keluarga akan berguna bagi keluarga dan pemerintah untuk membantu dirinya dalam menuntaskan keluarga dari ketertinggalan dan meningkatkan kualitas keluarga berdasarkan data yang telah terkelompokkan.

Indikator Keluarga Sejahtera pada dasarnya berangkat dari pokok pikiran yang terkandung di dalam undang-undang no. 10 Tahun 1992 disertai asumsi bahwa kesejahteraan merupakan variable komposit yang terdiri dari berbagai indikator yang spesifik dan operasional. Karena indikator yang yang dipilihakan digunakan oleh kader di desa, yang pada umumnya tingkat pendidikannya relative rendah, untuk mengukur derajat kesejahteraan para anggotanya dan sekaligus sebagai pegangan untuk melakukan melakukan intervensi, maka indicator tersebut selain harus memiliki validitas yang tinggi, juga dirancang sedemikian rupa, sehingga cukup sederhana dan secara

operasional dapat dipahami dan dilakukan oleh masyarakat di desa (Badan Keluarga Berencana Naional, 2012).

Pengklasifikasian keluarga yang dilakukan oleh para kader atau petugas setempat (Pos yandu kecamatan Bumiaji, Kota Batu) masih dilakukan secara manual, sehingga dirasa masih menyulitkan petugas. Petugas harus mempertimbangkan 21 nilai indikator yang telah diperoleh dari masing-masing keluarga. Jumlah keluarga dalam suatu wilayah desa atau kelurahan tidaklah sedikit, tentunya hal ini akan memakan waktu yang cukup lama bila proses pengklasifikasian dilakukan secara manual (Badan Keluarga Berencana Naional, 2012). Ketentuan dari instansi yang dijadikan acuan untuk proses pengklasifikasian biasanya masih dianggap rancu atau terkadang belum sesuai menurut perasaan si petugas yang bersangkutan, sehingga menimbulkan interpretasi tersendiri untuk mengklasifikasikannya. Sebenarnya sudah ada aplikasi online khusus dari pemerintah yang digunakan untuk manajemen pemutakhiran data keluarga yang beralamat pada http://aplikasi.bkkbn.go.id/, namun aplikasi ini belum mendukung fitur pengklasifikasian otomatis.

Oleh karena itu, perlu dilakukan penelitian agar menghasilkan suatu aplikasi yang dapat memudahkan dan mempercepat pengklasifikasian keluarga berdasarkan status kesejahteraannya sesuai dengan indicator yang ada. Salah satu metode yang dapat digunakan adalah Naive Bayes Classifier. Naïve Bayes Classifier adalah metode klasifikasi yang berdasarkan probabilitas dan Teorema Bayes dengan asumsi bahwa setiap variable X bersifat bebas, dengan kata lain, Naïve Bayes Classifier mengasumsikan bahwa keberadaan sebuah atribut (variable) tidak ada kaitannya dengan keberadaan atribut (variable) yang lain. Keuntungan penggunaan NBC adalah, metode ini hanya membutuhkan jumlah data pelatihan (data training) yang kecil untuk menentukan estimasi parameter yang diperlukan dalam proses pengklasifikasian (Tan, 2006].

Berdasarkan latar belakang yang telah diuraikan di atas, maka judul yang diambil untuk skripsi ini adalah **"Implementasi Metode Naïve Bayes Untuk Klasifikasi Taraf Hidup Masyarakat Sejahtera Pada Kota Batu".**

1.2 Rumusan Masalah

Berdasarkan latar belakang masalah diatas, maka rumusan masalah pada skripsi ini adalah:

1. Bagaimana melakukan implementasi metode Naïve Bayes pada kasus klasifikasi Keluarga Sejahtera?
2. Bagaimana menyajikan data hasil klasifikasi agar dapat diakses masyarakat?

1.3 Batasan Masalah

Batasan masalah pada skripsi ini adalah:

1. Parameter atau indicator yang digunakan untuk pengklasifikasian tingkat keluarga sejahtera adalah indikator yang sudah ditetapkan oleh BPMPPKB (Badan Pemberdayaan Masyarakat Perempuan dan Keluarga Berencana).
2. Untuk mengatasi *missing value* pada data latih dan data uji, akan digunakan rata rata sebagai pengganti *missing value.*
3. Data yang digunakan bersumber dari data hasil pemutakhiran keluarga tahun 2013 di Desa/Kelurahan Tulungrejo, RT. 01, 02, 03 – RW. 04 dan RT. 01, 02, 03, 04, 05 – RW. 05, Kecamatan Bumi Aji, Kota Batu

1.4 Tujuan

Tujuan yang ingin dicapai dari penulisan skripsi ini adalah:

1. Mengetahui cara melakukan implementasi metode Naïve Bayes pada kasus klasifikasi Keluarga Sejahtera.
2. Mengetahui tingkat akurasi algoritma pembelajaran Naïve Bayes untuk klasifikasi Keluarga sejahtera.

1.5 Manfaat

Manfaat yang diperoleh dari pengerjaan skripsi ini adalah, pemerintah dapat menggunakan data hasil klasifikasi yang telah disajikan dalam bentuk peta, sebagai pedoman dalam melakukan pemerataan pembanguanan, dari segi perekonomian, pendidikan, maupun kesehatan.

1.6 Sistematika Penulisan

Pembuatan tugas akhir ini dilakukan dengan sistematika sebagai berikut:

1. BAB I PENDAHULUAN

Bab ini berisi latar belakang penelitian, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, metodologi penelitian, dan sistematika penulisan.

2. BAB II TINJAUAN PUSTAKA

Bab ini berisi tentang teori dari berbagai pustaka yang menunjang untuk penelitian ini. Teori yang digunakan dalam penelitian ini antara lain mengenai keluarga sejahtera, google maps dan metode klasifikasi menggunakan Naive Bayes Classifier.

3. BAB III METODOLOGI DAN PERANCANGAN

Bab ini berisi tentang perancangan sistem perangkat lunak yang digunakan dalam penelitian ini, meliputi analisis data, analisis sistem, rencangan sistem, dan contoh perhitungan manual.

4. **BAB IV IMPLEMENTASI**

Berisi tentang segala hal yang berkaitan dengan implementasi sistem perangkat lunak yang digunakan untuk penelitian, meliputi implementasi source code dan antar muka Implementasi Metode Naïve Bayes Untuk Klasifikasi Taraf Hidup Masyarakat Sejahtera Pada Kota Batu

5. **BAB V PENGUJIAN DAN ANALISIS**

Berisi tentang penjelasn proses pengujian dan hasil pengujian serta analisis dari pengujian tersebut.

6. **PENUTUP**

Berisi kesimpulan yang diperoleh dari hasil pengujian dan saran untuk pengembangan.

BAB II. KAJIAN PUSTAKA DAN DASAR TEORI

Bab ini berisi kajian pustaka dan pembahasan mengenai teori dasar yang berhubungan dengan pengklasifikasian Keluarga Sejahtera menggunakan metode Naïve Bayes Classifier. Kajian pustaka membahas penelitian sebelumnya dan metode pada kasus yang berbeda. Dasar teori membahas teori penunjang yang berkaitan dengan penelitian diantaranya, Keluarga Sejahtera, Indikator Tahapan Keluarga Sejahtera, Klasifikasi, dan Naïve Bayes Classifier.

2.1 Kajian Pustaka

Berikut merupakan beberapa kajian pustaka yang digunakan penulis berdasarkan penelitian sebelumnya tentang permasalahan yang diangkat. Terdapat beberapa riset yang telah dilakukan untuk metode-metode dalam pengklasifikasian suatu objek. Beberapa penelitian sebelumnya adalah penelitian yang dilakukan oleh Devi dengan judul "Penentuan Mutu Buah Jeruk Manis Varietas Pacitan Berdasarkan Warna RGB Dan Diameter Dengan Metode Naïve Bayes Classifier". I Dw Gede A dengan judul "Klasifikasi Genre Pada Lagu Dengan Metode Naïve Bayes".

Penelitian pertama dengan judul "Penentuan Mutu Buah Jeruk Manis Varietas Pacitan Berdasarkan Warna RGB Dan Diameter Dengan Metode Naïve Bayes Classfier", dimana pada penelitian ini digunakan fitur RGB dan Diameter serta menggunakan metode Naïve Bayes Classifier, tujuan dari penelitian ini adalah mengklasifikasikan buah jeruk manis varietas pacitan sesuai dengan mutunya. Jeruk manis varietas pacitan akan diklasifikasikan terhadap 3 kriteria mutu, yaitu mutu A, B, dan C. Dengan menggunakan metode Naïve Bayes Classifier serta menggunakan fitur R, G, B dan Diameter, penelitian ini mendapatkan tingkat hasil akurasi sebesar 86% (Dewi, 2008).

Penelitian kedua dengan judul "Klasifikasi Genre Pada Lagu Dengan Metode Naïve Bayes", pada penelitian ini setiap data akan diekstraksi menjadi fitur. Hasil ekstraksi fitur tadi akan menghasilkan fitur berupa Frekuensi maksimal, median, dan minimal. Fitur yang dihasilkan pada awalnya berbentuk data continue dimana berisi angka numeric setiap fiturnya. Kemudian data fitur – fitur tersebut akan dikategorikan pada setiap fiturnya menjadi 3 bagian, dengan cara menentukan range data tersebut, dengan cara dikelompokkan menjadi frekuensi rendah, sedang, dan tinggi. Sehingga ketiga fiturnya merupakan fitur kategorikal. Aplikasi akan menghitung peluang setiap genre untuk lagu yang diinputkan menggunakan algoritma Naïve Bayes, dan kemudian dicari nilai yang paling maksimal pada setiap probabilitas kelas genre yang telah ditentukan, yakni punk, rock, slow, pop, dan dangdut. Pada penelitian ini didapatkan akurasi 60% disebabkan karena setiap lagi memilki frekuensi yang tidak menentu (Tan, 2006).

Data yang kita dapat dari ekstraksi fitur, terkadang tidak sesuai dengan apa yang kita prediksi diatas kertas. Pada kenyataanya selalu ada noise pada data set

dari proses ekstraksi fitur, dikarenakan adanya batasan batasan pada proses ekstraksi, yang menyebabkan data set yang kita terima, tidak mutlak seperti apa yang kita harapkan (Han, 2001). Apalagi saat kita mengambil banyak sampel untuk kita ekstraksi fiturnya, maka bisa jadi noise yang kita dapatkan akan semakin banyak. Metode Naïve Bayes tidak dapat memproses suatu fitur jika ditemukan padanya missing value, yaitu kekosongan nilai pada salah satu datanya. Akan tetapi karena Metode Naïve Bayes merupakan salah satu metode yang tingkat akurasinya tinggi, ketika suatu klasifikasi yang menggunakan algoritma Naïve Bayes ini menemukan missing value, maka menurut beberapa literatur, ketika menemukan missing value, kita dapat mengisi sendiri missing value tersebut dengan menggunakan nilai rata rata dari kelas tersebut (Caruana,2006).

2.2 Keluarga Sejahtera

Keluarga sejahtera adalah keluarga yang dibentuk berdasarkan atas perkawinan yang sah, mampu memenuhi kebutuhan hidup spiritual dan materiil yang layak, bertakwa kepada Tuhan Yang Maha Esa, memiliki hubungan yang serasi, selaras, dan seimbang antar anggota dan antar keluarga dengan masyarakat dan lingkungan (Undang – Undang Republik Indonesia nomor 52 tahun 2009).

Tingkat kesejahteraan keluarga dikelompokkan menjadi 5 (lima) tahapan, yaitu:

1. **Tahapan Keluarga Pra Sejahtera (KPS)**

   Yaitu keluarga yang tidak memenuhi salah satu dari 6 (enam) indikator Keluarga Sejahtera I (KS I) atau indikator “kebutuhan dasar keluarga” (basic needs).

2. **Tahapan Keluarga Sejahtera I (KS I)**

   Yaitu keluarga mampu memenuhi 6 (enam) indikator tahapan KS I, tetapi tidak memenuhi salah satu dari 8 (delapan) indikator Keluarga Sejahtera II atau indikator ” kebutuhan psikologis ” (psychological needs) keluarga.

3. **Tahapan Keluarga Sejahtera II**

   Yaitu keluarga yang mampu memenuhi 6 (enam) indikator tahapan KS I dan 8 (delapan) indikator KS II, tetapi tidak memenuhi salah satu dari 5 (lima) indikator Keluarga Sejahtera III (KS III), atau indikator ” kebutuhan pengembangan” (develomental needs) dari keluarga.

4. **Tahapan Keluarga Sejahtera III**

   Yaitu keluarga yang mampu memenuhi 6 (enam) indikator tahapan KS I, 8 (delapan) indikator KS II, dan 5 (lima) indikator KS III, tetapi tidak memenuhi

salah satu dari 2 (dua) indikator Keluarga Sejahtera III Plus (KS III Plus) atau indikator ”aktualisasi diri” (self esteem) keluarga.

5. Tahapan Keluarga Sejahtera III Plus

Yaitu keluarga yang mampu memenuhi keseluruhan dari 6 (enam) indikator tahapan KS I, 8 (delapan) indikator KS II, 5 (lima) indikator KS III, serta 2 (dua) indikator tahapan KS III Plus.

2.3 Indikator Tahapan Keluarga Sejahtera

A. Enam Indikator tahapan Keluarga Sejahtera I (KS I) atau indikator ”kebutuhan dasar keluarga” (basic needs), dari 21 indikator keluarga sejahtera yaitu:
   1. Pada umumnya anggota keluarga makan dua kali sehari atau lebih.
   2. Anggota keluarga memiliki pakaian yang berbeda untuk di rumah, bekerja/sekolah dan bepergian.
   3. Rumah yang ditempati keluarga mempunyai atap, lantai dan dinding yang baik.
   4. Bila ada anggota keluarga sakit dibawa ke sarana kesehatan.
   5. Bila pasangan usia subur ingin ber KB pergi ke sarana pelayanan kontrasepsi.
   6. Semua anak umur 7-15 tahun dalam keluarga bersekolah.

B. Delapan indikator Keluarga Sejahtera II (KS II) atau indikator ”kebutuhan psikologis” (psychological needs) keluarga, dari 21 indikator keluarga sejahtera yaitu:
   1. Pada umumnya anggota keluarga melaksanakan ibadah sesuai dengan agama dan kepercayaan masing-masing.
   2. Paling kurang sekali seminggu seluruh anggota keluarga makan daging/ikan/telur.
   3. Seluruh anggota keluarga memperoleh paling kurang satu stel pakaian baru dalam setahun.
   4. Luas lantai rumah paling kurang 8 m2 untuk setiap penghuni rumah.
   5. Tiga bulan terakhir keluarga dalam keadaan sehat sehingga dapat melaksanakan tugas/fungsi masing-masing.
   6. Ada seorang atau lebih anggota keluarga yang bekerja untuk memperoleh penghasilan.
   7. Seluruh anggota keluarga umur 10 - 60 tahun bisa baca tulisan latin.
   8. Pasangan usia subur dengan anak dua atau lebih menggunakan alat/obat kontrasepsi.

C. Lima indikator Keluarga Sejahtera III (KS III) atau indikator ”kebutuhan pengembangan ” (develomental needs), dari 21 indikator keluarga sejahtera yaitu:

1. Keluarga berupaya meningkatkan pengetahuan agama.
2. Sebagian penghasilan keluarga ditabung dalam bentuk uang atau barang.
3. Kebiasaan keluarga makan bersama paling kurang seminggu sekali dimanfaatkan untuk berkomunikasi.
4. Keluarga ikut dalam kegiatan masyarakat di lingkungan tempat tinggal.
5. Keluarga memperoleh informasi dari surat kabar/majalah/ radio/tv/internet.

D. Dua indikator Kelarga Sejahtera III Plus (KS III Plus) atau indikator ”aktualisasi diri” (self esteem) dari 21 indikator keluarga, yaitu:

1. Keluarga secara teratur dengan suka rela memberikan sumbangan materiil untuk kegiatan sosial.
2. Ada anggota keluarga yang aktif sebagai pengurBus perkumpulan sosial/yayasan/ institusi masyarakat.

2.4 Klasifikasi

Klasifikasi adalah suatu proses untuk mengelompokkan sejumlah data kedalam kelas – kelas tertentu yang sudah diberikan berdasarkan kesamaan sifat dan pola yang terdapat dalam data – data tersebut. Proses klasifikasi dimulai dengan diberikannya sejumlah data yang menjadi acuan untuk membuat aturan klasifikasi data. Data – data ini biasa disebut dengan training set. Dari training sets tersebut kemudian dibuat suatu model untuk mengklasifikasikan data. Model tersebut kemudian digunakan sebagai acuan untuk mengklasifikasikan data – data yang belum diketahui kelasnya yang biasa disebut dengan test sets. Beberapa metode klasifikasi adalah dengan menggunakan pohon keputusan, memory based reasoning, neural network, Naïve Bayes, dan support vector machine (Baktiar, 2013).

2.5 Naïve Bayes Classifier

Naïve Bayes merupakan teknik prediksi berbasis probabilitas sederhana yang berdasarkan pada penerapan teorema Bayes (aturan Bayes) dengan asumsi independen yang kuat (naïf) dimana diasumsukan kondisi antar fitur saling bebas, atau dengan kata lain metode ini mengasumsikan bahwa keberadaan sebuah fitur tidak ada kaitannya dengan keberadaan atribut yang lain (Rish, 2001). Jika X adalah vector masukan fitur dalam model klasifikasi dan Y adalah label kelas yang ada pada model klasifikasi, makan Naïve Bayes dapat ditulis dengan P(Y|X). Notasi tersebut berarti probabilitas label kelas Y didapat setelah fitur fitur X diamati. Formulasi Naïve Bayes untuk klasifikasi dirumuskan dalam persamaan (2-12) sebagai berikut:

$$P(Y \mid X) = \frac{P(Y)\prod_{i=1}^{q} P(X_i \mid Y)}{P(X)} \quad \text{..... (1)}$$

Dimana:

P(Y|X) adalah probabilitas data dengan vector X pada kelas Y

P(Y) adalah probabilitas awal kelas Y

$\prod_{i=1}^{q} P(X_{i|} \mid Y)$ adalah probabilitas independen kelas Y dari semua fitur dalam vector X

P(X) adalah probabilitas awal vector X

Nilai P(X) selalu tetap sehingga pada perhitungan prediksi nantinya tidak perlu dicantumkan. Sementara probabilitas independen $\prod_{i=1}^{q} P(X_{i|} \mid Y)$ merupakan pengaruh dari semua fitur dari data terhadap setiap kelas Y pada setiap set fitur $X = [X_1, X_2, X_3, ..., X_q]$ yang terdiri atas q atribut (q dimensi). Karena data yang didapat merupakan data kategorikal, maka kita menggunakan metode Naïve Bayes dengan rumus seperti pada persamaan 1.

$$h_{NaiveBayes} = \arg\max_h P(h)P(x \mid h) = \arg\max_h P(h)\prod_t P(a_t \mid h) \quad \text{..... (2)}$$

Keterangan:

P = probabilitas

h = hipotesis suatu class spesifik

P(h) = probabilitas hipotesis h

x = klasifikasi datum baru berupa x = ($a_{1,...}a_T$)

P(x|h) = probabilitas x berdasarkan kondisi pada hipotesis h

a_i = nilai fitur a_t dari setiap contoh datum

$P(a_{i|} \mid h) =$ probabilitas a_i berdasarkan kondisi pada hipotesis h

Untuk menghitung nilai varian (σ^2) maka rumus yang digunakan seperti persamaan 2.

$$\sigma^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2 \quad \text{..... (3)}$$

Keterangan:

- N = banyak datum yang dihitung
- x_i = datum ke i
- $\bar{x}$ = mean dari seluruh data

Berikut adalah algoritma dari Naïve Bayes:

Tentukan Data Latih dan Data Testing, cari mean, cari nilai varians, cari nilai densitas, cari posteriornya dengan rumus Naïve Bayes.

Varians merupakan kuadrat dari *standard deviation*, *standard deviation* atau sampangan baku merupakan ukuran penyebaran yang paling sering digunakan .Standar deviasi dari sebuah himpunan data adalah ukuran seberapa tersebarnya nilai data-data tersebut. Apabila penyebaran data sangat besar terhadap nilai rata-rata, maka nilai standar deviasi akan besar. Akan tetapi jika penyebaran data sangat kecil terhadap nilai rata-rata, maka nilai standar deviasi akan kecil. Persamaan standar deviasi pada suatu sampel data ditunjukkan pada persamaan sebagai berikut.

$$s_x = \sqrt{\frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}} \quad \text{..... (4)}$$

Dimana:

x_i = sampel data ke i

$\bar{X}$ = rata rata dari sampel data

$n =$ banyaknya data x pada suatu sampel

Berdasarkan persamaan diatas dapat diambil kesimpulan bahwa untuk mendapatkan nilai varians dari suatu sample dapat menggunakan persamaan sebagai berikut.

$$s_x^2 = \frac{\sum_{i=1}^{n}(x_i - \mu)^2}{n-1} \quad \text{..... (5)}$$

Contoh hitungan

Diberikan sebuah data uji berupa hewan musang dengan nilai fitur penutup kulit = rambut, melahirkan = ya, berat = 15. Masuk ke kelas manakan hewan musang tersebut, diberikan beberapa data training sebagai data latih untuk proses klasifikasi, dimana pada data training terdapat 2 kelas yaitu kelas mamalia dan kelas reptil, kemudian masing masing kels memiliki 3 fitur yaitu penutup kulit, melahirkan, dan berat badan. Contoh data training dapat dilihat pada Tabel 1.

Tabel 1. Contoh Data Training

Nama Hewan	Penutup Kulit	Melahirkan	Berat	Kelas
Ular	Sisik	Ya	10	Reptil
Tikus	Bulu	Ya	0.8	Mamalia
Kambing	Rambut	Ya	21	Mamalia
Sapi	Rambut	Ya	120	Mamalia

Nama Hewan	Penutup Kulit	Melahirkan	Berat	Kelas
Kadal	Sisik	Tidak	0.4	Reptil
Kucing	Rambut	Ya	1.5	Mamalia
Bekicot	Cangkang	Tidak	0.3	Reptil
Harimau	Rambut	Ya	43	Mamalia
Rusa	Rambut	Ya	45	Mamalia
Kura-Kura	Cangkang	Tidak	7	Reptil

Penyelesaian:

Hitung probabilitas setiap fitur pada setiap kelasnya atau $P(X_i|Y_j)$.

Perhitungan untuk fitur dengan tipe numerik adalah sebagai berikut:

Perhitungan untuk rata rata fitur berat badan pada kelas mamalia menggunakan persamaan (2-2) adalah sebagai berikut:

$$\mu_{mamalia} = \frac{0.8 + 21 + 120 + 1.5 + 43 + 45}{6} = \frac{231.3}{6} = 38.55$$

Perhitungan untuk rata rata fitur berat badan pada kelas reptil menggunakan persamaan (2-2) adalah sebagai berikut:

$$\mu_{reptil} = \frac{10 + 0.4 + 0.3 + 7}{4} = \frac{17.7}{4} = 4.425$$

Perhitungan untuk varians fitur berat badan pada kelas mamalia menggunakan persamaan (2-15) adalah sebagai berikut:

$$s_{mamalia}{}^2 = \frac{\begin{matrix}(0.8 - 38.55)^2 + (21 - 38.55)^2 + (120 - 38.55)^2 \\ + (1.5 - 38.55)^2 + (48 - 38.55)^2 + (45 - 38.55)^2\end{matrix}}{6 - 1}$$

$$s_{mamalia}{}^2 = \frac{9801.275}{5} = 1960.225$$

Perhitungan untuk varians fitur berat badan pada kelas mamalia menggunakan persamaan adalah sebagai berikut:

$$s_{reptil}{}^2 = \frac{(10 - 4.425)^2 + (0.4 - 4.425)^2 + (0.3 - 4.425)^2 + (7 - 4.425)^2}{4 - 1}$$

$$s_{reptil}{}^2 = \frac{70.9275}{3} = 23.6425$$

Perhitungan untuk peluang untuk fitur berat badan = 15 didalam kelas Mamalia dan reptil menggunakan persamaan (2-13) adalah sebagai berikut:

$$P(Berat = 15|Mamalia) = \frac{1}{\sqrt{2\pi 1960.225}} exp^{\frac{-(15-38.55)^2}{2\times 1960.225}} = 0.0104$$

$$P(Berat = 15|Reptil) = \frac{1}{\sqrt{2\pi 23.6425}} exp^{\frac{-(15-4.425)^2}{2\times 23.6425}} = 0.8733$$

Perhitungan untuk peluang untuk fitur kulit didalam kelas Mamalia adalah sebagai berikut:

P(Kulit=Sisik|Mamalia)=0/6
P(Kulit=Bulu|Mamalia)=1/6
P(Kulit=Rambut|Mamalia)=5/6
P(Kulit=Cangkang|Mamalia)=0/6

Perhitungan untuk peluang untuk fitur kulit didalam kelas Reptil adalah sebagai berikut:

P(Kulit=Sisik|Reptil)=2/4
P(Kulit=Bulu|Reptil)=0/4
P(Kulit=Rambut|Reptil)=0/4
P(Kulit=Cangkang|Reptil)=2/4

Perhitungan untuk peluang untuk fitur Melahirkan didalam kelas Reptil adalah sebagai berikut:

P(Lahir=Ya|Mamalia)=6/6
P(Lahir=Tidak|Mamalia)=0/6

Perhitungan untuk peluang untuk fitur Melahirkan didalam kelas Reptil adalah sebagai berikut:

P(Lahir=Ya|Reptil)=1/4
P(Lahir=Tidak|Reptil)=3/4

Hasil perhitungan dapat dilihat pada Tabel

Tabel 2. Hasil Perhitungan Probabilitas Setiap Fitur

Fitur Penutup Kulit		Fitur Melahirkan	
Mamalia	Reptil	Mamalia	Reptil
Sisik = 0 Bulu = 1 Rambut = 5 Cangkang = 0 Total Mamalia = 6	Sisik = 2 Bulu = 0 Rambut= 0 Cangkang = 2 Total Reptil = 4	Ya = 6 Tidak = 0	Ya = 1 Tidak = 3
P(Kulit=Sisik|Mamalia)=0 P(Kulit=Bulu|Mamalia)=1/6 P(Kulit=Rambut|Mamalia)=5/6 P(Kulit=Cangkang|Mamalia)=0	P(Kulit=Sisik|Reptil)=0.5 P(Kulit=Bulu|Reptil)=0 P(Kulit=Rambut|Reptil)=0 P(Kulit=Cangkang|Reptil)=0.5	P(Lahir=Ya|Mamalia)=1 P(Lahir=Tidak|Mamalia)=0	P(Lahir=Ya|Reptil)=0.25 P(Lahir=Tidak|Reptil)=0.75

Fitur Berat		Kelas	
Mamalia	Reptil	Mamalia	Reptil
$\mu_{mamalia} = 38.55$ $\sigma_{mamalia}{}^2 = 1960.255$	$\mu_{mamalia} = 38.55$ $\sigma_{mamalia}{}^2 = 23.6425$	Mamalia=6 P(Mamalia)= 6/10	Reptil=4 P(Reptil)=4/10

Menghitung probabilitas akhir untuk setiap kelas sesuai dengan hasil perhitungan pada tabel 2.3, persamaan (2-21) dan persamaan (2-22):

$$P(X|Mamalia) = P(Kulit = Rambut|Mamalia) \times P(Lahir = Ya|Mamalia)$$

$$\times P(Berat = 15|Mamalia) = \frac{5}{6} \times 1 \times 0.0104 = 0.0087$$

$$P(X|Reptil) = P(Kulit = Rambut|Reptil) \times P(Lahir = Ya|Reptil) \times$$

$$P(Berat = 15|Reptil) = 0 \times 0.25 \times 0.8733 = 0$$

Menghitung probabilitas akhir sesuai dengan hasil perhitungan pada tabel 2.3, persamaan (2-23) dan persamaan (2-24):

$$P(Mamalia|X) = 0.6 \times 0.0087 = 0.0052$$

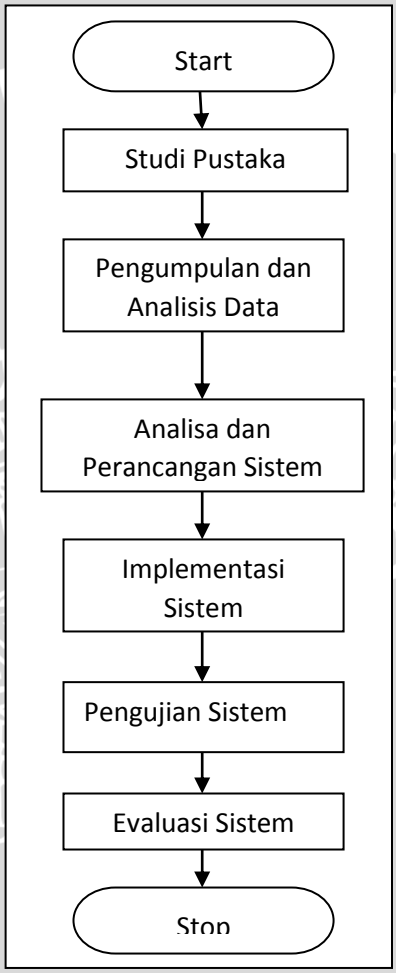
$$P(Reptil|X) = 0.4 \times 0 = 0$$

Karena nilai probabilitas akhir terbesar ada di kelas mamalia dengan nilai 0.0052 sehingga data uji musang diprediksi sebagai kelas mamalia.

BAB III. METODOLOGI DAN PERANCANGAN

3.1. Metodologi

Bab ini menjelaskan mengenai algoritma dan langkah – langkah yang akan digunakan dalam penelitian Implementasi Metode *Naive Bayes* Untuk Klasifikasi Taraf Hidup Masyarakat Sejahtera Pada Kecamatan Batu. Penjelasann tersebut merupakan tujuan untuk memberikan gambaran mengenai sistem yang akan dibuat berdasarkan tinjauan pustaka pada bab sebelumnya. Gambar 3.1 menjelaskan mekanisme penelitian secara umum.



Gambar 1. Diagram Alir Penelitian

Penjelasan dari diagram blog diatas adalah sebagai berikut:

1. Studi literatur berkaitan dengan data survei Kesejahteraan Masyarakat Kecamatan Batu, Naive Bayes
2. Pengumpulan data survei kesejahteraan masyarakat kecamatan Batu yang nantinya akan digunakan penentuan tingkat kesejahteraan masyarakat secara otomatis
3. Perancangan sistem untuk pengklasifikasian otomatis dengan metode Naive Bayes pada data survei yang ada.
4. Implementasi sistem berdasarkan perancangan an analisis yang telah dilakukan berkaitan dengan klasifikasi otomatis dengan metode Naive Bayes.
5. Pengujian dan analisis terhadap sistem yang telah diimplementasikan.
6. Penarikan kesimpulan dari penelitian yang telah dilakukan.

3.2 Perancangan

3.2.1 Analisis Data

Data yang digunakan dalam skripsi ini didapatkan dari Posyandu Kecamatan Bumiaji, Kota Batu. Data yang diambil adalah data indikator Keluarga Sejahtera, dan status TKS pada data pemutakhiran keluarga tahun 2013 di Desa/Kelurahan Tulungrejo, RT 01, 02, 03 – RW 04 dan RT 01, 02, 03, 04, 05 – RW 05, Kecamatan Bumiaji, Kota Batu. Contoh data ditampilkan pada tabel 3.1.

Tabel 3.1 Contoh Data Indikator dan Status TKS

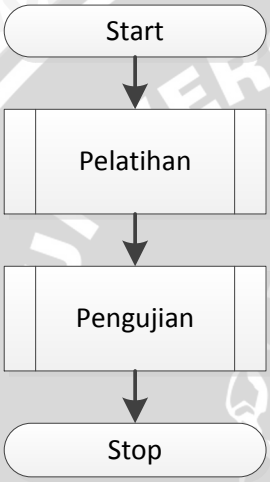
Indikator	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	
IN																						
7	v	v	v	v	v	v	v	v	v	v	v	v	v	v	v	v	v	v	v	v	v	KSIII+
10	v	v	v	v	v	v	v	v	v	v	v	v	v	v	v	v	v	v	v	v	v	KSIII+
1	v	v	v	v	v	-	v	v	v	v	v	v	v	v	v	v	v	v	v	x	x	KSIII
4	v	v	v	v	v	-	v	v	v	v	v	v	v	-	v	v	v	v	v	x	x	KSIII
2	v	v	v	v	v	-	v	v	v	v	v	v	v	v	v	v	x	v	v	x	x	KSII
5	v	v	v	v	-	v	v	v	v	v	v	v	v	v	v	v	x	v	v	x	x	KSII
3	v	v	v	v	v	v	v	v	v	x	v	v	v	v	v	x	v	v	v	x	x	KSI
122	v	v	v	v	v	-	v	v	v	x	v	v	v	v	v	x	v	v	v	x	x	KSI
6	v	v	x	v	v	v	v	v	v	v	v	v	v	v	v	x	v	v	v	x	x	KPS
55	v	v	v	v	-	x	v	v	v	v	v	v	v	-	x	x	v	v	v	x	x	KPS

Data tersebut adalah data indikator yang menentukan Tahapan Keluarga Sejahtera (TKS tiap keluarga. Indikator yang digunakan ada 21 dengan 5 macam status TKS. Nilai yang terdapat pada tiap inikator berbeda – beda, terdapat 3 macam nilai dari 21 indikator yaitu apabila responden memiliki salah satu indikator yang ditentukan, maka pada form akan ditandai dengan simbol (v), bila responded tidak memiliki indikator tersebut, maka akan di tandai dengan (x), dan ada beberapa error pada pengumpulan data, yaitu indikator yang kosong atau

tidak diisi, sehingga tidak diketahui apakah responded tersebut memiliki indikator tersebut atau tidak disimbolkan dengan ( ) data tersebut dinamakan Missing Value.

3.2.2 Analisis Proses

Pada Bagian ini dijelaskan rancangan proses serta gambaran umum bagaimana proses kerja sistem membentuk suatu model probabilistik dan pengklasifikasian menggunakan metode Naive Bayes. Proses yang pertama adalah Pembacaan data testing, pembuatan model probabilistik. Sedang proses yang kedua adalah pengujian data testing dan pengklasifikasian dengan menggunakan metode Naive Bayes. Tahapan yang terjadi pada proses pertama dan kedua akan digambarkan dengan flowchart dibawah ini

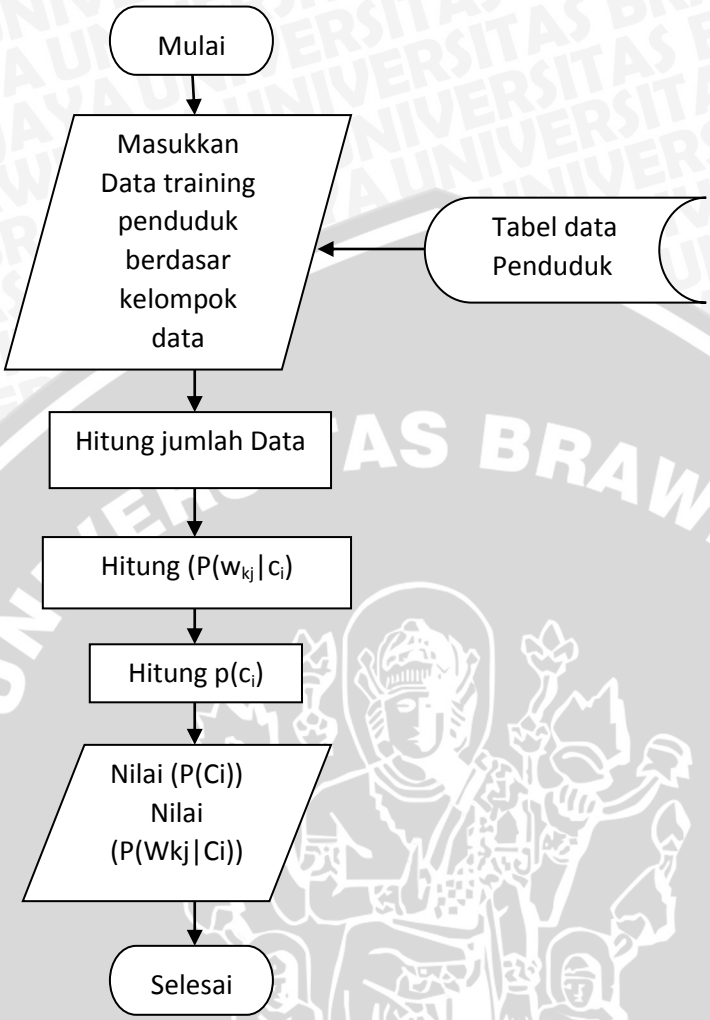


Gambar 2 Flowchart Sistem

Gambaran umum proses proses utama yang akan dilakukan oleh sistem dalam penelitian ini dijelaskan pada Gambar 3.2.

Detail dari masing masing proses pada Gambar 3.2 akan dijelaskan lebih lanjut pada sub bab ini mulai dari proses data training perhitungan probabilistik hingga penentuan Tingkat Keluarga Sejahtera.

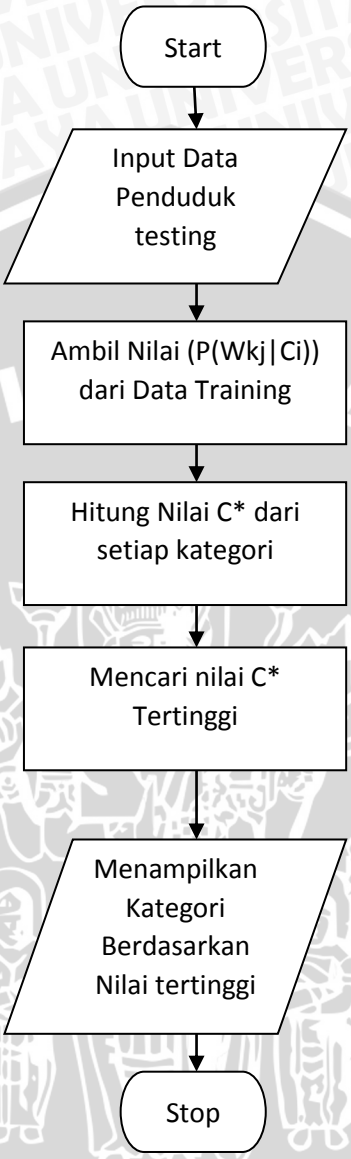
3.2.2.1 Data Training



Gambar 3. Flowchart Data Training

Proses data training ini dimulai dari penggolongan data berdasar Tingkat Keluarga Sejahtera, selanjutnya dilakukan perhitungan untuk menentukan model probabilstik dari sample data yang sudah disediakan.

3.2.2.2. Data Testing



Gambar 4. Flowart Data Testing

Pada bagian ini, dimasukkan data yang akan diujikan, kemudian dari data yang akan diuji tersebut cari nilai (P(Wkj|Ci)) kemudian hitung nilai C* dari setiap kategori, dari hasil yang di dapat, akan didapat nilai terbesar yang memiliki kesamaan dari salah satu kategori

3.2.3 Hitungan Manual

Dalam pembentukan aturan klasifikasi keluarga sejahtera yang menggunakan metode naive bayes, diberikan contoh perhitungan manual untuk gambaran awal system dalam melakukan pembentukan aturan klasifikasi, hingga

pengecekan hasil akhir untuk memastikan system telah bekerja sesuai dengan algoritma yang benar.

Tahap I

Data yang akan digunakan untuk perhitungan manual padatahap I terdiri dari 10 data awal pada data indicator dan status TKS pada data hasil pemutakhiran keluarga tahun 2013 di Desa/Kelurahan Tulungrejo, RT. 01, 02, 03, 04, 05 – RW. 05, Kecamatan Bumiaji, Kota Batu.

Tabel 4. Data Indicator

Indikator	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	
7	v	v	v	v	v	v	v	v	v	v	v	v	v	v	v	v	v	v	v	v	v	KSIII+
10	v	v	v	v	v	v	v	v	v	v	v	v	v	v	v	v	v	v	v	v	v	KSIII+
1	v	v	v	v	v	-	v	v	v	v	v	v	v	v	v	v	v	v	v	x	x	KSIII
4	v	v	v	v	v	-	v	v	v	v	v	v	v	-	v	v	v	v	v	x	x	KSIII
2	v	v	v	v	v	-	v	v	v	v	v	v	v	v	v	v	x	v	v	x	x	KSII
5	v	v	v	v	-	v	v	v	v	v	v	v	v	v	v	v	x	v	v	x	x	KSII
3	v	v	v	v	v	v	v	v	v	x	v	v	v	v	v	x	v	v	v	x	x	KSI
122	v	v	v	v	v	-	v	v	v	x	v	v	v	v	v	x	v	v	v	x	x	KSI
6	v	v	x	v	v	v	v	v	v	v	v	v	v	v	v	x	v	v	v	x	x	KPS
55	v	v	v	v	-	x	v	v	v	v	v	v	v	-	x	x	v	v	v	x	x	KPS

Pertama tama adalah mengelompokkan sejumlah data sesuai dengan kategori sebagai data training, data yang dipilih sebaiknya data yang memiliki error paling sedikit, error disini yang dimaksud adalah missing value atau data yang hilang, error pada data ini mempengaruhi akurasi dari hasil permodelan nantinya. Selanjutnya data data yang dipilih tersebut dikelompokkan sesuai dengan kategorinya. Semakin banyak data yang sama dalam satu kategori akan meningkatkan akurasi dari permodelan klasifikasinya. Proses penentuan kategori dari sebuah dokumen testing dilakukan perhitungan dengan menggunakan metode berikut

$$c* = \arg \max_{ci \in C} p(ci|dj)$$

$$= \arg \max_{ci \in C} \prod p(Wkj|C*) \times p(Ci) \quad \text{.......................................(8)}$$

Dimana Wkj merupakan fitur atau indikator dari data survey. Kumpulan data yang dipilih akan dibentuk term document matrix. Langkah selanjutnya adalah pembuatan model probabilistik dengan melakukan perhitungan

$$p(Wkj|Ci) = \frac{f(Wkj, Ci) + 1}{f(Ci) + |W|} \quad \text{.....(9)}$$

$f(Wkj, ci)$ adalah nilai kemunculan kata Wkj pada kategori ci

$f(ci)$ adalah jumlah keseluruhan kata pada kategori ci

$|W|$ adalah jumlah keseluruhan kata/fitur yang digunakan

Contoh penyelesaian perhitungan:

Data contoh perhitungan diambil dari tabel 3, dengan data keluarga nomor urut 7 pada indikator pertama

$$p(Wkj|Ci) = \frac{f(Wkj, Ci) + 1}{f(Ci) + |W|}$$

$$p(Wkj|Ci) = \frac{2 + 1}{42 + 21}$$

$$p(Wkj|Ci) = \frac{3}{63}$$

$$p(Wkj|Ci) = 0.04761$$

Hasil dari perhitungan sebesar 0.04761 yang selanjutnya proses yang sama akan diulangi untuk setiap fitur dari setiap kategori yang digunakan sebagai data latih. Sehingga nantinya akan didapatkan model probabilistic dari masing - masing data yang akan digunakan untuk melakukan perbandingan dengan data testing.

Tabel 5. Data Training

Kategori	P Kategori	IND-1	2	3	4	5	6	7	IND-20	IND-21
KSIII+	0.2	0.047	0.047	0.047	0.047	0.047	0.047	0.047		0.047619
KSIII	0.2	0.053	0.053	0.053	0.053	0.053	0.017	0.053		0.017857
KSII	0.2	0.054	0.054	0.054	0.054	0.036	0.036	0.054		0.018181
KSI	0.2	0.055	0.055	0.055	0.055	0.055	0.037	0.055		0.018515
KPS	0.2	0.056	0.056	0.037	0.056	0.037	0.037	0.056		0.018867

Setelah semua data set yang dipilih telah dirubah menjadi tebel model probabilistik, maka selanjutnya dipilih sample data yang nantinya akan digunakan sebagai data uji. Dari seluruh data survey dipilih beberapa data secara acak, tidak terikat berapapun jumlahnya. Sebagai contoh pada perhitungan manual kali ini

digunakan 5 data random sebagai data yang nantinya akan digunakan untuk menguji klasifikasi dari program yang dibuat sudah sesuai dengan tujuannya.

3.3. Perancangan Uji Coba

Perancangan uji coba akurasi 1 menggunakan rancangan uji coba tabel 3.1. Proses uji coba pertama dilakukan sebanyak 3 kali dengan menggunakan jumlah data latih yang berbeda dengan nilai k yang berbeda. Data latih pertama adalah sebesar 80%dari data, data latih kedua sebesar 70% dari data, data latih ketiga sebesar 60% dari data. Sedangkan data uji adalah data yang belum digunakan dalam proses pelatihan yaitu sebesar 20% dari data, sebesar 30% dari data, dan sebesar 40% dari data. Perancangan skenario uji coba sistem ditunjukkan pada tabel berikut:

Tabel 6. Rancangan Uji Coba 1

Nilai K	Nilai Akurasi (%)			Akurasi Rata-rata
	Data Latih 1	Data Latih 2	Data Latih 3	
2				
..				
21				

Perancangan uji coba akurasi 2 menggunakan rancangan tabel 6 Proses uji coba kedua dilakukan sebanyak 3 kali dengan menggunakan jumlah data latih yang berbeda dengan nilai k terbaik yang didapatkan dari uji coba 1. Data latih pertama adalah sebesar 10 data, data latih kedua sebesar 15 data, data latih ketiga sebesar 20 data. Sedangkan data uji adalah data yang belum digunakan dalam proses pelatihan yaitu sebesar 40 data. Perancangan skenario uji coba sistem ditunjukkan pada tabel berikut:

Tabel 7. Rancangan Uji Coba 2

Data	Akurasi
Data Latih 1	
Data Latih 2	
Data Latih 3	

Perancangan uji coba akurasi 3 menggunakan rancangan tabel 7. Proses uji coba ketiga dilakukan sebanyak 10 kali dengan menggunakan jumlah data latih terbaik pada uji coba kedua dengan nilai k terbaik yang didapatkan dari uji coba pertama. Data latih yang digunakan adalah sebesar 50 data. Sedangkan data uji adalah data yang belum digunakan dalam proses pelatihan yaitu sebesar 40 data. Rancangan uji coba yang ketiga ini adalah uji coba akurasi dengan data seimbang

dan tidak seimbang. Perancangan skenario uji coba sistem ditunjukkan pada tabel berikut:

Tabel 8. Rancangan Uji Coba 3

Uji Coba Ke	Data Latih Seimbang	Data Latih Tidak Seimbang
1		
...		
10		

3.4 Perancangan Antarmuka

Tampilan Awal

Gambar 5. Halaman Utama

BAB IV. IMPLEMENTASI

4.1 Lingkungan Implementasi

Lingkungan implementasi yang akan dijelaskan dalam sub bab ini adalah lingkungan implementasi perangkat lunak dan perangkat keras yang digunakan dalam mengimplementasikan sistem yang telah dibuat dalam penelitian ini.

4.1.1 Lingkungan Perangkat Keras

Perangkat keras yang digunakan dalam mengembangkan perangkat lunak dalam penelitian ini memiliki spesifikasi sebagai berikut:

1. Intel(R) Core(TM) i3-2370M CPU @ 2.40GHz (4 CPUs)
2. Memori 6 GB
3. Harddisk 500 GB
4. Monitor 14”

4.1.2 Lingkungan Perangkat Lunak

Perangkat lunak yang digunakan dalam mengembangkan sistem dan penelitian ini terdiri dari:

1. Sistem Operasi Windows 7 Home Basic 64bit
2. JDK 7.45 64-bit
3. NetBeans IDE 7.1

4.2 Implementasi Program

Implementasi program akan menjelaskan implementasi prosedur – prosedur dari perancangan yang telah dirancang sebelumnya kedalam system, yang mana menggunakan bahasa pemrograman JAVA. Secara garis besar system dibangun dari 3 modul yaitu, membuat model probalilistik, membuat data training, dan data testing. Pada aplikasinya system yang di buat terdiri dari 3 Package yaitu image yang berfungsi sebagi package yang menampung file gambar yang digunakan pada system, MyGUI adalah package yang berisi class class desain dari system yang di buat, dan mymatrixone yang berisi class class yang berfungsi sebagai proses dari system ini, mulai dari pembentukan dataset, Stemming, Data training, hingga Pengujian data. Berikut akan dijelaskan lebih lanjut mengenai proses-proses penting yang dijalankan oleh system.

4.2.1 Load Document

Dalam penelitian ini system membaca data yang berasal dari sebuah file excel yang berisi data indicator yang sudah ditentukan. Pembacaan data dilakukan dengan membaca data yang ada pada file excel dan menyusun formatnya dengan membuat array yang digunakan untuk membentuk tabel. Menampilkan data yang akan di proses setelah dibaca dari file .xls dan disajikan dalam bentuk tabel.

```
private void
jButton_load_docActionPerformed(java.awt.event.ActionEvent evt) {
        // LOAD DOKUMEN
        String excelXLSFileName = "datadamar.xls";
        vectorDataExcelXLS = readDataExcelXLS(excelXLSFileName);
        byk_baris=vectorDataExcelXLS.size()-1;
        byk_kolom=23;


        Object[] namaKolom = {"Data Keluarga ke-i", "Indikator 1",
"Indikator 2", "Indikator 3", "Indikator 4", "Indikator 5", "Indikator 6"
        , "Indikator 7", "Indikator 8", "Indikator 9", "Indikator 10",
"Indikator 11", "Indikator 12", "Indikator 13", "Indikator 14", "Indikator
15", "Indikator 16"
        , "Indikator 17", "Indikator 18", "Indikator 19", "Indikator 20",
"Indikator 21", "Kelas"};


        Object[][] obj_data_doc= new Object[byk_baris][byk_kolom];
        obj_data= new Object[byk_baris][byk_kolom-1];
        for(int i=0;i<byk_baris;i++){
           obj_data_doc[i][0]="Doc ke-".concat(Integer.toString(i+1));
        }

        for(int kolom = 0; kolom < byk_kolom-1;kolom++){

getSpecifyvectorDataExcelXLS=getSpecifyColumnOfReadDataExcelXLS(v
ectorDataExcelXLS,kolom);
          //System.out.println("");
          for(int i=0;i<byk_baris;i++){
             obj_data[i][kolom]=getSpecifyvectorDataExcelXLS.get(i);
             obj_data_doc[i][kolom+1]=obj_data[i][kolom];
          }
        }
        DefaultTableModel dtm_obj= new
DefaultTableModel(obj_data_doc, namaKolom);
        jTable4.setModel(dtm_obj);
        jTable4.setAutoResizeMode(JTable.AUTO_RESIZE_OFF);
        jTable3.setAutoResizeMode(JTable.AUTO_RESIZE_OFF);
        setTitle("Load Document");
        setVisible(true);
    }
```

4.2.2. Preprocessing

Pada bagian preprocessing ini system akan melakukan penyeleksian, dan pembersihan terhadap data data yang tidak ada nilainya, karena dari data mentah terdapat cukup banyak noise yang berupa missing data yang dimana hal ini tidak dapat diterima oleh system, karena itu nilai nilai kosong tersebut akan dihapuskan.

```
private void
jButton_preproActionPerformed(java.awt.event.ActionEvent evt) {
      Object[] namaKolom = {"Data Keluarga ke-i", "Indikator 1",
"Indikator 2", "Indikator 3", "Indikator 4", "Indikator 5", "Indikator 6"
      , "Indikator 7", "Indikator 8", "Indikator 9", "Indikator 10",
"Indikator 11", "Indikator 12", "Indikator 13", "Indikator 14", "Indikator
15", "Indikator 16"
      , "Indikator 17", "Indikator 18", "Indikator 19", "Indikator 20",
"Indikator 21", "Kelas"};


      Object[][] obj_data_doc= new Object[byk_baris][byk_kolom];
      for(int i=0;i<byk_baris;i++){
         obj_data_doc[i][0]="Doc ke-".concat(Integer.toString(i+1));
      }


      for(int kolom = 0; kolom < byk_kolom-1;kolom++){
         //System.out.println("");
         for(int i=0;i<byk_baris;i++){
            if(obj_data[i][kolom]==""){
               obj_data[i][kolom]=MissingValue(obj_data, i, kolom);
               obj_data_doc[i][kolom+1]=obj_data[i][kolom];
            }
            else{
               obj_data_doc[i][kolom+1]=obj_data[i][kolom];
            }
         }
      }


      DefaultTableModel dtm_obj= new
DefaultTableModel(obj_data_doc, namaKolom);

      jTable4.setModel(dtm_obj);

      jTable4.setAutoResizeMode(JTable.AUTO_RESIZE_OFF);

      setTitle("Load Document");

      setVisible(true);
```

```
    }

    private Object ChangeFormat(Object[][] obj,int x,int y){
        Object hasil="";
        Object test = "v";
        Object test1 = "x";
        if(obj[x][y].equals(test)){
            hasil =1;
        }
        else if(obj[x][y].equals(test1)){
            hasil =0;
        }
        else{
            hasil = (obj[x][y]);
        }
        return hasil;
    }
```

4.2.3 Konversi Data

Pada bagian ini terdapat fungsi yang akan mengkonversi check list (v) ataupun tanda x menjadi angka 1 dan 0 yang dapat di proses oleh system

```
private void
jButton_create_datasetActionPerformed(java.awt.event.ActionEvent
evt) {

    jTabbedPane1.setSelectedIndex(1);

    // TODO add your handling code here:

    Object[] namaKolom = {"Indikator 1", "Indikator 2", "Indikator 3",
"Indikator 4", "Indikator 5", "Indikator 6"

    , "Indikator 7", "Indikator 8", "Indikator 9", "Indikator 10",
"Indikator 11", "Indikator 12", "Indikator 13", "Indikator 14", "Indikator
15", "Indikator 16"

    , "Indikator 17", "Indikator 18", "Indikator 19", "Indikator 20",
"Indikator 21", "Kelas"};

    Object[][] obj_data_doc= new Object[byk_baris][byk_kolom];

    for(int kolom = 0; kolom < byk_kolom-1;kolom++){

      //System.out.println("");
```

```
        for(int i=0;i<byk_baris;i++){
            obj_data[i][kolom]=ChangeFormat(obj_data, i, kolom);
            obj_data_doc[i][kolom]=obj_data[i][kolom];
        }
    }
    DefaultTableModel dtm_obj= new
DefaultTableModel(obj_data_doc, namaKolom);
    jTable1.setModel(dtm_obj);
    jTable1.setAutoResizeMode(JTable.AUTO_RESIZE_OFF);
    setTitle("Create Data");
    setVisible(true);
  }
```

4.2.4 Data Training

Pada bagian ini dilakukan proses perhitungan, mulai dari penentuan nilai peluang setiap kelas, memecah dokumen perkelas, dan menentukan model probabilitas untuk setiap indicator per kelas.

```
private void jButton1ActionPerformed(java.awt.event.ActionEvent evt) {
    jTabbedPane1.setSelectedIndex(2);
    // TODO add your handling code here:
    Object[] namaKolom = {"Kelas","Probabilitas Kelas", "Indikator 1",
"Indikator 2", "Indikator 3", "Indikator 4", "Indikator 5", "Indikator 6"
    , "Indikator 7", "Indikator 8", "Indikator 9", "Indikator 10",
"Indikator 11", "Indikator 12", "Indikator 13", "Indikator 14", "Indikator
15", "Indikator 16"
    , "Indikator 17", "Indikator 18", "Indikator 19", "Indikator 20",
"Indikator 21"};
    Object[][] obj_data_doc= new Object[5][byk_kolom];
    // nama kelas2nya
```

```
        obj_data_doc[0][0]="KPS";
        obj_data_doc[1][0]="KSI";
        obj_data_doc[2][0]="KSII";
        obj_data_doc[3][0]="KSIII";
        obj_data_doc[4][0]="KSIII+";
        //probabilitas kelas
        obj_data_doc[0][1]="0.2";
        obj_data_doc[1][1]="0.2";
        obj_data_doc[2][1]="0.2";
        obj_data_doc[3][1]="0.2";
        obj_data_doc[4][1]="0.2";

        //memecah dokumen per kelas
        double[][] TrainingDT = Naive.TrainingDT(obj_data, byk_baris);
        for (int i = 0; i < 5; i++) {
           for (int j = 2; j < namaKolom.length; j++) {
              obj_data_doc[i][j]= TrainingDT[i][j-2];
              obj_data[i][j-2] = obj_data_doc[i][j];
           }
        }
        DefaultTableModel dtm_obj= new
DefaultTableModel(obj_data_doc, namaKolom);
        jTable2.setModel(dtm_obj);
        jTable2.setAutoResizeMode(JTable.AUTO_RESIZE_OFF);
        setTitle("Training Data");
        setVisible(true);
    }
```

4.2.5 Naïve Bayes

Dibawah ini merupakan implementasi dari naïve bayes

```
public class NaiveB {
  public static int totalterm [];
  public double[][] TrainingDT(Object [][]vsm,int byk_baris){
     double[][] Hasil;
     int Nkelas= 5;
     //pecah data
     ArrayList [] dttrain = PecahData(vsm, byk_baris);
     //hitung total bilangan c tiap kelas
     int [][] NC = GetNC(dttrain, byk_baris);
     //hitung total fitur
     int Nfitur = 21;
     //hitung total data pada tiap kelas
     int [] NW = GetNW(NC);
     //hitung PC Kelas
     Hasil = GetPC(NC, Nfitur, NW);
     return Hasil;
  }
  public ArrayList [] PecahData (Object[][] obj_data,int byk_baris){
    ArrayList<Object[]> [] Hasil = new ArrayList[5];
    for(int i = 0 ; i<5; i++){
       ArrayList<Object[]> obj_data_Hasil = new ArrayList<>();
       Hasil[i]=obj_data_Hasil;
    }
    for(int i = 0 ; i<byk_baris;i++){
       if(obj_data[i][21].equals("kps")||obj_data[i][21].equals("KPS")){
          Hasil[0].add(obj_data[i]);
       }
       else
if(obj_data[i][21].equals("ksi")||obj_data[i][21].equals("KSI")){
          Hasil[1].add(obj_data[i]);
       }
       else
if(obj_data[i][21].equals("ksii")||obj_data[i][21].equals("KSII")){
          Hasil[2].add(obj_data[i]);
       }
       else
if(obj_data[i][21].equals("ksiii")||obj_data[i][21].equals("KSIII")){
          Hasil[3].add(obj_data[i]);
       }
       else
if(obj_data[i][21].equals("ksiii+")||obj_data[i][21].equals("KSIII+")){
```

```
            Hasil[4].add(obj_data[i]);
          }
          else{
            System.out.println("gagal memilah objek"+obj_data[i][21]);
          }
        }
      return Hasil;
    }

    public int [][] GetNC (ArrayList<Object[]> [] dttrain,int byk_baris){
      int [][] hasil = new int [5][21];
      for(int i = 0 ; i <5 ;i++){
        for (int j = 0; j < dttrain[i].size(); j++) {
          for (int k = 0; k < 21; k++) {
            if(dttrain[i].get(j)[k].equals(1)){
              hasil[i][k]++;
            }
          }
        }
      }
      return hasil;
    }

     public int [] GetNW (int [][] NC){
       int [] hasil = new int [5];
         for (int i = 0; i < NC.length; i++) {
           for (int j = 0; j < NC[i].length; j++) {
             hasil[i]=hasil[i]+NC[i][j];
           }
         }
       return hasil;
    }

    public double [][] GetPC (int [][] NC,int Nfitur,int [] NW){
      double [][] hasil = new double [5][21];
      DecimalFormat df = new
DecimalFormat(".####################");

      for (int i = 0 ; i<NC.length;i++) {
        for (int j = 0; j < NC[1].length; j++) {
          double a =NC[i][j]+1;
          double b = NW[i]+Nfitur;
          //hasil[i][j]=(NC[i][j]/(NW[i]+Nfitur)*1000000000);
          hasil[i][j]=a/b;
          df.format(hasil[i][j]);
```

```
        }
      }
       return hasil;
   }
```

4.2.6 Data Testing

Data testing merupakan inti dari implementasi system ini, dimana akan dilakukan pengujian, apakah system telah melakukan klasifikasi secara otomatis dan membandingkan akurasi system dengan kondisi yang sebenarnya. Dengan data latih yang telah disiapkan sebelumnya, maka bisa dimasukkan secara manual data dari indicator yan ingin diujikan. Selajutnya akan keluar hasil perhitungan dari setiap kelas, dan system akan otomatis memilih nilai tertinggi karena nilai tersebut adalah nilai yang paling mendekati.

```
private void jButton2ActionPerformed(java.awt.event.ActionEvent evt) {
    jTabbedPane1.setSelectedIndex(3);
  }
   private Object [] Testing(String dttest,int index) throws
FileNotFoundException, IOException{
      String dttesString = dttest;
      String data [][];
      Stemming a = new Stemming();
      VSpace b = new VSpace();
      data = b.addDT(dttesString);
      Object namaKolom[] = new Object[data.length+1];
      //insert nama kolom
      for(int i = 0; i < data.length;i++){
         if(i==0){
            namaKolom[i]=" ";
         }
         else{
         namaKolom[i]=data[i][0];
         }
      }
      Object namaKolom1[] = new Object[2];
      //insert nama kolom
      for(int i = 0; i < 2;i++){
         if(i==0){
            namaKolom1[i]="kelas";
         }
         else{
         namaKolom1[i]="probabilitas";
         }
```

```
    }
    Object obj_dataset_random1 [] = new Object[7];
    double datapc [] = Naive.GetPC(hasilprob);
    //ambil hasil klasifikasi
    Object max [][]= Naive.GetKategori(datapc);
    String klsasli = GetKelas(index);
    boolean hasil = (max[0][0].equals(klsasli));
    for(int i=0;i<obj_dataset_random1.length;i++){
      if (i==5){
        obj_dataset_random1[i]=max[0][0];
      }
      else if(i==6){
        obj_dataset_random1[i]=hasil+"";
      }
      else{
        obj_dataset_random1[i]=datapc[i];
      }
    }
    return obj_dataset_random1;
  }
```

BAB V. PENGUJIAN

Bab ini membahas tentang tahapan pengujian dan analisis hasil dari system Klasifikasi Keluarga Sejahtera dengan *Naïve Bayes*. Pengujian yang akan dilakukan mengevaluasi hasil uji coba sesuai dengan perancangan uji coba sebelumnya dan mengevaluasi akurasi dari system Klasifikasi Keluarga Sejahtera dengan metode *Naïve Bayes.*

5.1. Hasil dan analisis akurasi hasil sistem

Untuk mengetahui akurasi dari system yang dibuat, dibutuhkan suatu pengujian untuk mengetahui seberapa besar akurasi dari system Klasifikasi Keluarga sejahtera menggunakan metode *Naïve Bayes.* Ada berbagai alternatif pengujian yang dilakukan, tetapi pada dasarnya Naïve Bayes adalah suatu metode yang tidak memperhatikan jarak antar data, sehingga pengujian untuk mengetahui factor factor yang berhubungan ketepatan program dalam menuliskan hasil akan dibahas nanti. Disini akan dibahas mengenai perbandingan output dari system dan akan dibandingan dengan hasil asli, dimana nantinya akan diketahui akurasi keseluruhan dari program.

Skenario pengujian pertama adalah memberikan data latih sebanyak 100 data pada system, yang selanjutnya hasil pengujian akan ditampilkan pada table dengan penjelasan Pengujian yaitu urutan pengujian yang telah di lakukan. No adalah id atau nomor dari keluarga yang digunakan dalam pengujian. Akurasi ketika pengujian telah dilakukan, setelah diketahui hasilnya, apabila sesuai dengan data sebenarnya, maka ditandai dengan angka 1, sedangkan apabila tidak sesuai, diberi tanda angka 0, yang nantinya memudahkan untuk membuat persentase akurasi dari system.

Tabel 9. Pengujian 100 Data Uji

Pengujian	id	Akurasi
1	2	1
2	4	1
3	6	1
4	8	1
5	10	1
6	12	0
7	14	1
8	16	1
9	18	1
10	20	1
11	22	1
12	24	1
13	26	1
14	28	1

Pengujian	id	Akurasi
15	30	1
16	32	1
17	34	1
18	36	1
19	38	0
20	40	1
21	42	1
22	44	1
23	46	0
24	48	1
25	50	1
26	52	1
27	54	1
28	56	1

Pengujian	id	Akurasi
29	58	1
30	60	1
31	62	1
32	64	1
33	66	1
34	68	1
35	70	1
36	72	1
37	74	1
38	76	1
39	78	1
40	80	1
41	82	1
42	84	1

Pengujian	id	Akurasi
43	86	1
44	88	1
45	90	1
46	92	1
47	94	1
48	96	1
49	98	1
50	100	0
51	7	1
52	101	1
53	214	1
54	195	1
55	203	1

Pada pengujian scenario pertama ini, didapatkan hasil sebagai berikut. Dari 55 kali pengujian, data yang berhasil di klasifikasi dengan tepat sebanyak 51 data dari 55, dengan demikian akurasi yang di dapat adalah 92,73%.

Skenario pengujian Kedua adalah memberikan data latih sebanyak 20 data pada system, yang selanjutnya hasil pengujian akan ditampilkan pada table dengan penjelasan Pengujian yaitu urutan pengujian yang telah di lakukan. No adalah id atau nomor dari keluarga yang digunakan dalam pengujian. Akurasi ketika pengujian telah dilakukan, setelah diketahui hasilnya, apabila sesuai dengan data sebenarnya, maka ditandai dengan angka 1, sedangkan apabila tidak sesuai, diberi tanda angka 0, yang nantinya memudahkan untuk membuat persentase akurasi dari system.

Tabel 10. Pengujian dengan 20 Data Latih

Pengujian ke	Nomer id	Hasil
1	2	0
2	4	1
3	6	1
4	8	1
5	10	0
6	12	1
7	14	1
8	16	0
9	18	1
10	20	1

Pengujian ke	Nomer id	Hasil
11	22	1
12	24	0
13	26	0
14	28	0
15	30	1
16	32	1
17	34	1
18	36	1
19	38	0
20	40	1

Pengujian ke	Nomer id	Hasil
21	42	1
22	44	1
23	46	0
24	48	0
25	50	1
26	52	1
27	54	0
28	56	1
29	58	1
30	60	1
31	62	1
32	64	1
33	66	0
34	68	1
35	70	1
36	72	0
37	74	1
38	76	1
39	78	1
40	80	1
41	82	1
42	84	1
43	86	0
44	88	1
45	90	1
46	92	1
47	94	1
48	96	0
49	98	1
50	100	1

Pada pengujian scenario kedua ini, didapatkan hasil sebagai berikut. Dari 50 kali pengujian, data yang berhasil di klasifikasi dengan tepat sebanyak 36 data dari 55, dengan demikian akurasi yang di dapat adalah 72%.

Tabel 11. Pengujian dengan 10 Data Latih

Pengujian ke	Nomer id	Hasil
1	2	0
2	4	1
3	6	1
4	8	1
5	10	1
6	12	0
7	14	0
8	16	0
9	18	0
10	20	1
11	22	0
12	24	1
13	26	0
14	28	0
15	30	1
16	32	0
17	34	0
18	36	1
19	38	1
20	40	1
21	42	1
22	44	1
23	46	0
24	48	0
25	50	0
26	52	1
27	54	0
28	56	1
29	58	0
30	60	1
31	62	0
32	64	1
33	66	1
34	68	1

Pengujian ke	Nomer id	Hasil
35	70	0
36	72	1
37	74	1
38	76	0
39	78	1
40	80	1
41	82	1
42	84	0
43	86	1
44	88	1
45	90	0
46	92	1
47	94	1
48	96	0
49	98	0
50	100	1
51	102	1
52	104	1
53	106	1
54	108	1
55	110	1

Pada pengujian scenario ketiga, didapatkan hasil sebagai berikut. Dari 55 kali pengujian, data yang berhasil di klasifikasi dengan tepat sebanyak 36 data dari 55, dengan demikian akurasi yang di dapat adalah 60%.

Tabel 12. Hasil Uji Coba Pada Akurasi Klasifikasi Keluarga Sejahtera

Skenario	Banyak Data Latih	Pengujian	Akurasi
1	10	55x	60%
2	20	50x	72,00%
3	100	55x	92,73%

Dari tabel percobaan di atas dapat disimpulkan bahwa, semakin banyak data latih yang diberikan untuk program, semakin tinggi pula akurasinya. Untuk variasi jumlah data pengujian, semakin banyak pengujian yang dilakukan, semakin presisi pula akurasi yang di dapatkan.

BAB VI. PENUTUP

6.1. Kesimpulan

Berdasarkan penelitian dengan judul Implementasi Metode Naïve Bayes Untuk Klasifikasi Taraf Hidup Masyarakat Sejahtera Pada Kecamatan Batu dapat di ambil kesimpulan bahwa:

- Hasil uji coba skenario pertama menunjukkan bahwa 100 data latih setelah dilakukan pengujian sebanyak 55 kali mendapatkan akurasi 92,73%
- Hasil uji coba skenario kedua 20 data latih, dilakukan 50 pengujian, didapatkan akurasi 72%
- Hasil uji coba skenario ketiga, dengan 10 data latih, dan dilakukan 55 kali pengujian, didapatkan akurasi sebesar 60%

6.2. Saran

Berdasarkan penelitian judul Implementasi Metode Naïve Bayes Untuk Klasifikasi Taraf Hidup Masyarakat Sejahtera Pada Kecamatan Batu dapat di sarankan bahwa kedepannya kita penelitian ini dilanjutkan, akan digunakan metode lain yang dapat memproses data missing value, sehingga akan didapatkan hasil klasifikasi yang lebih akurat. Proses ujicoba yang dilakukan agar lebih bervariasi agar didapatkan hasil yang optimal

DAFTAR PUSTAKA

Baktiar, Yoga Agung. 2013. “Implementasi Metode Naive Bayes Untuk Klasifikasi Kenaikan Grade Karyawanpada Fuzzyfikasi Data Kinerja Karyawan”. Skripsi Program Teknologi Informasi Dan Ilmu Komputer Uniersitas Brawijaya : Malang.

Badan Kependudukan dan Keluarga Berencana Nasional. 2012. “Profil Hasil Pendataan Keluarga Tahun 2012”, Direktorat Pelaporan dan Statistik, Jakarta.

Caruana, R.; Niculescu-Mizil, A. 2006. An Empirical Comparison of Supervised Learning Algorithms. Proc. 23rd International Conference on Machine Learning. CiteSeerX 10.1.1.122.5901

Dewi, Evi R. 2008. Penentuan Mutu Buah Jeruk Manis Varietas Pacitan Berdasarkan Warna RGB Dan Diameter Dengan Metode Naie Bayes Classifier. Skripsi FMIPA Universitas Brawijaya : malang.

Gosling, James; McGilton, Henry (May 1996). "The Java Language Environment".

Han, Jiawei dan Micheline Khamber. 2001. “Data Mining: Concepts and Technique”, Morgan Kaufmann Publisher, San Francisco, USA.

Hand, D. J.; Yu, K. 2001. "Idiot's Bayes — not so stupid after all?". International Statistical Review. 69 (3): 385 - 399. doi:10.2307/ 1403452. ISSN 0306-7734

Ling, Bai. & Wei, Guo Zhu. 2006, *Journal of Cancer Molecules 2*, No.4, hal. 141-153, Guilin, China.

Mitchell, Tom. 1997. “Machine Learning”, Singapore, McGraw-Hill.

National Geographic Indonesia (in Indonesian). Hanya ada 13.466 Pulau di Indonesia. 8 February 2012.

A Conversation with James Gosling – ACM Queue". Queue.acm.org. 2004-08-31. Retrieved 2010-06-09.

Rish, Irina. 2001. An empirical study of the naive Bayes classifier. IJCAI Workshop on Empirical Methods in AI

Stuart J. Russell and Peter Norvig. 2003. Artificial Intelligence: A Modern Approach (2 ed.). Pearson Education. See p. 499 for reference to

"idiot Bayes" as well as the general definition of the Naive Bayes model and its independence assumptions

Tan P. N. 2006. “Introduction to Data Mining”, Addison Wesley.

Wahana, Jaka dan Kirbrandoko. 1995. Pengantar Mikro Ekonomi Jilid 1, Terjemahan Cetakan Pertama, Binaruksa Aksara, Jakarta