

SKRIPSI

KLASIFIKASI KELAYAKAN PENERIMA BANTUAN
PROGRAM KELUARGA HARAPAN MENGGUNAKAN
RANDOM FOREST
(Studi Kasus: Kelurahan Desa Balerejo)



ABDUL ROFIQ ALMUQOROBIN
NPM. 17.0504.0051

PROGRAM STUDI TEKNIK INFORMATIKA S1
FAKULTAS TEKNIK
UNIVERSITAS MUHAMMADIYAH MAGELANG

2021

SKRIPSI

KLASIFIKASI KELAYAKAN PENERIMA BANTUAN
PROGRAM KELUARGA HARAPAN MENGGUNAKAN
RANDOM FOREST
(Studi Kasus: Kelurahan Desa Balerejo)

Disusun Sebagai Salah Satu Syarat Memperoleh Gelar Sarjana Komputer
(S.Kom)

Program Studi Teknik Informatika Jenjang Strata Satu (S-1) Fakultas Teknik
Universitas Muhammadiyah Magelang



ABDUL ROFIQ ALMUQOROBIN
NPM. 17.0504.0051

PROGRAM STUDI TEKNIK INFORMATIKA S1
FAKULTAS TEKNIK
UNIVERSITAS MUHAMMADIYAH MAGELANG

2021

HALAMAN PENEGASAN

Skripsi ini adalah hasil karya sendiri, dan semua sumber baik yang dikutip dan dirujuk telah saya nyatakan dengan benar.

Nama : Abdul Rofiq Almuqorobin

NPM : 17.0504.0051

Magelang, 12 Agustus 2021



Abdul Rofiq Almuqorobin

17.0504.0051

PERNYATAAN KEASLIAN

Yang bertanda tangan di bawah ini:

Nama : Abdul Rofiq Almuqorobin
NPM : 17.0504.0051
Program Studi : Teknik Informatika S1
Fakultas : Teknik
Alamat : Krajan Balerejo RT/RW 05/01 Kec.Kaliangkrik
Kab.Magelang.
Judul Skripsi : Klasifikasi Kelayakan Penerima Bantuan Program
Keluarga Harapan Menggunakan *Random Forest*
(Studi Kasus: Kelurahan Desa Balerejo)

Dengan ini menyatakan bahwa Skripsi ini merupakan hasil karya sendiri dan bukan merupakan plagiat dari hasil karya orang lain. Dan bila di kemudian hari terbukti bahwa karya ini merupakan plagiat, maka saya bersedia menerima sanksi administrasi maupun sanksi apapun.

Demikian surat pernyataan ini saya buat dengan penuh kesadaran dan sebenarnya serta penuh tanggung jawab.

Magelang, 12 Agustus 2021

Yang menyatakan



Abdul Rofiq Almuqorobin

17.0504.0051

HALAMAN PENGESAHAN

SKRIPSI

KLASIFIKASI KELAYAKAN PENERIMA BANTUAN PROGRAM
KELUARGA HARAPAN MENGGUNAKAN *RANDOM FOREST*
(Studi Kasus: Kelurahan Desa Balerejo)

Disusun Oleh :

Abdul Rofiq Almuqorobin

NPM. 17.0504.0051

Telah dipertahankan di depan Dewan Penguji

Pada Tanggal 12 Agustus 2021

Susunan Dewan Penguji

Pembimbing I

Dr. Uky Yudatama, S.Si., M.Kom
NIDN. 0605107201

Pembimbing II

Maimunah, S. Si., M. Kom
NIDN. 0612117702

Penguji I

Mukhtar Hanafi, ST., M.Cs
NIDN. 0602047502

Penguji II

Endah Ratna Arumi, M.Cs.
NIDN. 0601129001

Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer

Tanggal, 12 Agustus 2021

Dekan



Yun Arifatul Fatimah, ST., MT., Ph.D
NIK. 987408139

HALAMAN PERNYATAAN PERETUJUAN

LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Yang bertanda tangan di bawah ini, saya:

Nama : Abdul Rofiq Almuqorobin
NPM : 17.0504.0051
Fakultas/ Jurusan : Teknik / Teknik Informatika S-1
E-mail address : abdulrofiqalmuqorobin@gmail.com

demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Perpustakaan UNIMMA, Hak Bebas *Royalty Non-Eksklusif (Non-exclusive Royalty-Free Right)* atas karya ilmiah LKP/ KP TA/ SKRIPSI TESIS Artikel Jurnal *) yang berjudul :

KLASIFIKASI KELAYAKAN PENERIMA BANTUAN PROGRAM KELUARGA HARAPAN MENGGUNAKAN *RANDOM FOREST* (Studi Kasus: Kelurahan Desa Balerejo)

beserta perangkat yang diperlukan (bila ada). Dengan Hak Bebas *Royalty Non-Eksklusif (Non-Exclusive Royalty-Free Right)* ini Perpustakaan UNIMMA berhak menyimpan, mengalih-media/format-kan, mengelolanya dalam bentuk pangkalan data (*database*), mendistribusikannya, dan menampilkan/ mempublikasikannya di internet atau media lain untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis/ pencipta dan atau penerbit yang bersangkutan.

Saya bersedia untuk menanggung secara pribadi, tanpa melibatkan pihak Perpustakaan UNIMMA, segala bentuk tuntutan hukum yang timbul atas pelanggaran Hak Cipta dalam karya ilmiah saya ini.

Demikian pernyataan ini dibuat dengan sesungguhnya.

Dibuat di : Magelang
Pada tanggal : 12 Agustus 2021

Penulis,



Abdul Rofiq Almuqorobin
NPM. 17.0504.0051

*) : *pilih salah Satu*

Mengetahui,
Dosen Pembimbing

Dr. Uky Yudatama, S.Si., M.Kom.
NIDN. 0605107101

KATA PENGANTAR

Puji syukur kehadiran Allah SWT, atas segala rahmat dan hidayah-Nya sehingga penulis dapat menyelesaikan skripsi ini dengan lancar. Skripsi ini disusun sebagai salah satu syarat untuk mencapai gelar Sarjana Komputer di Program Studi Teknik Informatika S1 Fakultas Teknik Universitas Muhammadiyah Magelang. Penyelesaian Skripsi ini banyak memperoleh bantuan dan bimbingan dari berbagai pihak. Oleh karena itu, diucapkan terima kasih kepada :

1. Dr. Lilik Andriyani, S.E., M.Si selaku Rektor Universitas Muhammadiyah Magelang.
2. Yun Arifatul Fatimah, S.T., M.T., Ph.D. selaku Dekan Fakultas Teknik Universitas Muhammadiyah Magelang
3. Endah Ratna Arumi, M.Cs selaku Ketua Program Studi Teknik Informatika S1 Universitas Muhammadiyah Magelang.
4. Dr. Uky Yudatama, S.Si., M.Kom dan Maimunah, S.Si., M.Kom selaku Dosen pembimbing, yang telah menyediakan waktu, tenaga, dan pikiran untuk mengarahkan penyusunan Skripsi ini.
5. Seluruh Dosen Fakultas Teknik Universitas Muhammadiyah Magelang yang telah memberikan ilmu dan pengetahuan yang bermanfaat.
6. Seluruh bagian kantor Desa Balerejo yang telah banyak membantu dalam usaha memperoleh data yang diperlukan.
7. Orang tua dan keluarga yang telah memberikan bantuan dukungan material dan moral.
8. Teman-Teman yang telah membantu dan memberikan dukungan baik secara moril maupun materil hingga terselesaikannya skripsi ini.

Akhir kata, semoga Allah SWT berkenan membalas segala kebaikan semua pihak yang telah membantu yang tidak dapat disebutkan namanya satu persatu dan semoga Skripsi ini membawa manfaat bagi semua pihak.

Magelang, 12 Agustus 2021



Abdul Rofiq Almuqorobin
NPM. 17.0504.0051

DAFTAR ISI

HALAMAN KULIT MUKA	i
HALAMAN JUDUL.....	ii
HALAMAN PENEGASAN.....	iii
PERNYATAAN KEASLIAN.....	iv
HALAMAN PENEGASAN.....	v
HALAMAN PERNYATAAN PERETUJUAN	vi
KATA PENGANTAR	vii
DAFTAR ISI.....	viii
DAFTAR TABEL.....	x
DAFTAR GAMBAR	xi
DAFTAR LAMPIRAN.....	xii
INTISAR.....	xiii
ABSTRACT.....	xiv
BAB 1 PENDAHULUAN	1
1.1 Latar Belakang Permasalahan.....	1
1.2 Rumusan Masalah	2
1.3 Tujuan Penelitian.....	3
1.4 Manfaat Penelitian.....	3
BAB 2 TINJAUAN PUSTAKA	4
2.1 Penelitian yang Relevan	4
2.2 Landasan Teori	7
2.2.1 Kemiskinan	7
2.2.2 Program Keluarga Harapan.....	7
2.2.3 <i>Data Mining</i>	8
2.2.4 Pengelompokan Data Mining.....	9
2.2.5 Tahap Tahab Data Mining	10
2.2.6 Klasifikasi	11
2.2.7 <i>Random Forest</i>	12
2.2.8 Python	14
2.2.9 Flask.....	14
2.2.10 <i>K-fold Cross Validation</i>	14

2.2.11	<i>Synthetic Minority Over-sampling Technique</i>	15
BAB 3	METODOLOGI PENELITIAN.....	16
3.1	Prosedur Penelitian.....	16
3.2	Metode Pengumpulan Data.....	21
3.3	Metode Pengolahan Data.....	22
BAB 5	KESIMPULAN DAN SARAN	38
5.1	Kesimpulan	38
5.2	Saran.....	38
DAFTAR PUSTAKA	39

DAFTAR TABEL

DAFTAR GAMBAR

Gambar 2. 1 KDD <i>Process</i> (Sumber : (Antara et al. 2015)).....	8
Gambar 2. 2 Bidang ilmu data mining (Sumber (Mardi 2017)).....	9
Gambar 2. 3 Tahab tahab data mining (Sumber (Fatmawati and Windarto 2018))	10
Gambar 2. 4 Ilustrasi metode <i>random forest</i> sumber (Iman and Wahyu 2021)....	12
Gambar 3. 1 Kerangka kerja penelitian.....	16
Gambar 3. 2 <i>confusion matrix</i> (Sumber (Setyawan and Suradi 2017))	18
Gambar 3. 3 Halaman Login.....	19
Gambar 3. 4 Halaman <i>Home</i>	19
Gambar 3. 5 Halaman Data.....	20
Gambar 3. 6 Halaman Klasifikasi	20
Gambar 3. 7 Halaman Hasil	21
Gambar 3. 8 Konsep sederhana <i>random forest</i> (Sumber(Danoedoro and Murti 2021)).....	22

DAFTAR LAMPIRAN

INTISAR

Klasifikasi Kelayakan Penerima Bantuan Program Keluarga Harapan Menggunakan *Random Forest* (Studi Kasus: Kelurahan Desa Balerejo)

Oleh : Abdul Rofiq Almuqorobin

Pembimbing : 1. Dr. Uky Yudatama, S.Si., M.Kom

2. Maimunah, S.Si., M.Kom.

Kemiskinan merupakan salah satu masalah yang dialami oleh beberapa negara berkembang, termasuk Indonesia. Dikarenakan banyaknya penduduk, lapangan pekerjaan terbatas, yang mengakibatkan banyak pengangguran. Dalam mengatasi permasalahan tersebut, pemerintah tentu tidak tinggal diam. Berbagai upaya telah dilakukan untuk mengatasi kemiskinan, salah satunya dengan adanya program PKH (Program Keluarga Harapan) yang berada dibawah koordinasi Tim Koordinasi Penanggulangan Kemiskinan (TKPK), baik dipusat maupun di daerah. Klasifikasi Kelayakan Penerima Bantuan Program Keluarga Harapan Menggunakan *Random Forest*. *Random forest* merupakan metode yang dapat meningkatkan hasil akurasi, karena dalam membangkitkan simpul anak untuk setiap *node* dilakukan secara acak. untuk melakukan evaluasi pada penelitian ini menggunakan *confusion matrix*. Penelitian ini menggunakan *random forest* terbaik adalah yang memiliki jumlah 5 pohon. data yang digunakan adalah data penduduk miskin yang terdapat pada desa balerejo yang berjumlah 503 kemudian data tersebut dibagi menjadi data *training* dan data *testing*. Dengan perbandingan 30% data *training* dan 70% data *testing* dan didapatkan hasil akurasi klasifikasi 74.1%.

Kata Kunci : *random forest*, klasifikasi, *data mining*, kemiskinan, program keluarga harapan

ABSTRACT

***Eligibility Classification Of Hope Family Program Recipients Using Random Forest
(Case Study: Balerejo Village)***

By : Abdul Rofiq Almuqorobin
Supervisor : 1. Dr. Uky Yudatama, S.Si., M.Kom
2. Maimunah, S.Si., M.Kom.

Poverty is one of the problems experienced by several developing countries, including Indonesia. Due to the large population, employment opportunities are limited, resulting in a lot of unemployment. In overcoming these problems, the government certainly does not remain silent. In overcoming these problems, the government certainly does not remain silent. Various efforts have been made to reduce poverty, one of which is the PKH program (Program Keluarga Harapan) which is under the coordination of the Poverty Reduction Coordination Team (TKPK), both at the center and in the regions. Eligibility Classification of Family Hope Program Assistance Beneficiaries Using Random Forest. Random forest is a method that can improve accuracy results, because the generation of child nodes for each node is done randomly. to evaluate in this study using a confusion matrix. This study uses a random forest, the best is the one that has 5 trees. The data used is the data of the poor population in Balerejo village, totaling 503 then the data is divided into training data and testing data. With a comparison of 30% of training data and 70% of testing data, the classification accuracy result is 74.1%.

Keywords: random forest, classification, data mining, poverty, hope family program

BAB 1

PENDAHULUAN

1.1 Latar Belakang Permasalahan

Kemiskinan merupakan salah satu masalah yang dialami oleh beberapa negara berkembang, termasuk Indonesia. Dikarenakan banyaknya penduduk, lapangan pekerjaan terbatas, yang mengakibatkan banyak pengangguran. Penggolongan kemiskinan didasarkan pada suatu standart tertentu yaitu dengan membandingkan tingkat pendapatan orang atau keluarga dengan tingkat pendapatan yang diperlukan untuk memenuhi kebutuhan pokok minimum. Dalam mengatasi permasalahan tersebut, pemerintah tentu tidak tinggal diam. Berbagai upaya telah dilakukan untuk mengatasi kemiskinan, salah satunya dengan adanya program PKH (Program Keluarga Harapan) yang berada dibawah koordinasi TimKoordinasi Penanggulangan Kemiskinan (TKPK), baik dipusat maupun didaerah. PKH adalah program perlindungan social yang memberikan bantuan tunai kepada Rumah Tangga Sangat Miskin (RTSM). Program ini dalam jangka pendek bertujuan mengurangi beban RTSM dan dalam jangka panjang diharapkan dapat memutus mata rantai kemiskinan antar generasi, sehingga generasi berikutnya dapat keluar dari perangkap kemiskinan (Anggraeni, Arifiana, and Abadi 2017).

Desa Balerejo merupakan salah satu wilayah yang berada di Kecamatan Kaliangkrik daerah barat yang memiliki jumlah RW (Rukun Warga) sebanyak 11 RW dimana di tiap RW tersebut memiliki beberapa RT (Rukun Tetangga) dengan jumlah penduduk kurang lebih 3 239 jiwa. Data statistik tahun 2019 dari badan pusat statistik Kabupaten Magelang (Sataloff, Johns, and Kost n.d.). Dari data tersebut tercatat jumlah warga miskin (gakin) sebanyak 503 Kepala Keluarga (KK). Dari data warga miskin tersebut, Penerima bantuan PKH 207 KK. Dengan adanya masalah tersebut, mengakibatkan adanya ketidaksinkronan penerima bantuan PKH. Masyarakat yang berhak menerima bisa jadi tidak menerima, sedangkan warga yang bukan merupakan warga miskin malah justru menerima bantuan PKH. jika penerima bantuan PKH jelas terdata dengan baik, serta syarat kelayakan penerimaan bantuan PKH yang terdata dengan baik, maka akan

mengurangi ketidaksinkronan penerima bantuan PKH dan juga lebih memudahkan petugas di dalam pembagian bantuan PKH tersebut (Sugianto and Maulana 2019).

Data mining adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Teknik-teknik, metode-metode, atau algoritma dalam *data mining* sangat bervariasi (Mardi 2017). Salah satu teknik *data mining* adalah fungsi klasifikasi. Fungsi klasifikasi memudahkan untuk mengidentifikasi kelayakan penerima dan bukan penerima bantuan PKH. Teknik klasifikasi adalah teknik pembelajaran yang digunakan untuk memprediksi nilai dari atribut kategori target (Sugianto and Maulana 2019). Salah satu metode klasifikasi yang bisa digunakan adalah *random forest*. *Random forest* merupakan metode yang dapat meningkatkan hasil akurasi, karena dalam membangkitkan simpul anak untuk setiap *node* dilakukan secara acak. Metode ini digunakan untuk membangun pohon keputusan yang terdiri dari *root node*, *internal node*, dan *leaf node* dengan mengambil atribut dan data secara acak sesuai ketentuan yang diberlakukan (Yusuf Sulisty Nugroho 2017) .

Berdasarkan uraian permasalahan diatas, dapat diketahui bahwa dalam mengatasi permasalahan ketidaksinkronan dalam penerima bantuan PKH pada Desa Balerejo. Oleh karena itu dalam penelitian ini akan dibangun sebuah sistem yang mampu melakukan klasifikasi terhadap masyarakat yang layak dan tidak layak menerima bantuan PKH. Dengan adanya klasifikasi data penduduk miskin agar penerima bantuan PKH bisa tepat sasaran. Sehingga diambil judul penelitian **“Klasifikasi Kelayakan Penerima Bantuan Program Keluarga Harapan Menggunakan Random Forest (Studi Kasus: Kelurahan Desa Balerejo)”**

1.2 Rumusan Masalah

Berdasarkan uraian latar belakang diatas, rumusan masalah yang harus diselesaikan dalam penelitian ini yaitu, Bagaimana menerapkan *random forest* dalam suatu sistem untuk mengklasifikasikan tingkat kemiskinan penerima bantuan PKH.

1.3 Tujuan Penelitian

Berdasarkan Rumusan masalah diatas, maka tujuan peneliti yang akan dicapai adalah:

1. Menentukan kelayakan keluarga penerima bantuan PKH.
2. Melakukan klasifikasi kelayakan penerima PKH menggunakan menggunakan *random forest*.

1.4 Manfaat Penelitian

Berdasarkan tujuan penelitian yang telah disebutkan di atas, maka hasil penelitian ini adalah:

1. Memberikan informasi dan masukan kepada pemerintah desa dalam menentukan layak tidaknya menerima bantuan PKH.
2. Mendapatkan hasil klasifikasi agar bisa digunakan untuk mengetahui layak tidaknya penerima PKH.

BAB 2

TINJAUAN PUSTAKA

2.1 Penelitian yang Relevan

Penelitian yang dilakukan (Yusuf Sulistyono Nugroho 2017) yang berjudul “Sistem Klasifikasi Variabel Tingkat Penerimaan Konsumen Terhadap Mobil Menggunakan Metode *Random Forest*” permasalahan yang di hadapi yaitu tingkat penjualan mobil di Indonesia masih belum stabil akibat tingkat penerimaan konsumen yang rendah. Ada banyak faktor yang mempengaruhi tingkat penerimaan konsumen terhadap produk produk mobil yang baru diluncurkan. Jika hal tersebut terus terjadi, akan berimbas buruk pada profitabilitas produsen mobil di negara ini. Sehingga dibutuhkan sebuah sistem untuk mengklasifikasikan faktor yang mempengaruhi permasalahan tingkat penerimaan mobil oleh pihak konsumen. Peneliti ini menggunakan metode *random forest* untuk melakukan klasifikasi terhadap tingkat penerimaan konsumen terhadap mobil. Hasil dari peneliti ini adalah sistem klasifikasi yang akan digunakan sebagai pertimbangan bagi produsen mobil dimasa mendatang dalam proses produksi mobil.

Penelitian yang dilakukan (Iman and Wahyu 2021) yang berjudul “Klasifikasi Rumah Tangga Penerima Beras Miskin (Raskin)/Beras Sejahtera (Rastra) di Provinsi Jawa Barat Tahun 2017 dengan Metode *Random Forest* dan *Support Vector Machine*” permasalahan yang dihadapi yaitu mengenai kualitas beras, jumlah beras yang diterima, terutama rumah tangga sasaran. Terdapat 17.5% rumah tangga di Indonesia memenuhi syarat pada tiga program bantuan sosial. Sayangnya masih banyak yang tidak menerima satupun bantuan sosial, termasuk raskin/rastra. Sehingga peneliti melakukan pendekatan *data mining* untuk mengelompokkan keluarga yang berhak menerima bantuan Raskin/Rastra diharapkan lebih tepat sasaran dan pemanfaatannya dapat semakin dirasakan masyarakat. Peneliti ini menggunakan metode *random forest* dan svm untuk melakukan klasifikasi terhadap rumah tangga penerima

beras miskin (Raskin)/ beras sejahtera (Rastra) di Provinsi Jawa Barat tahun 2017. Hasil evaluasi model yang dilakukan pada kedua metode, *random forest* dan SVM, ketelitian yang dihasilkan hampir sama. Selisih kinerja klasifikasi yang dihasilkan kedua metode *random forest* dan SVM tidak signifikan oleh karena itu, dapat dikatakan bahwa kedua metode klasifikasi memiliki kinerja yang setara dan sama baiknya dalam melakukan klasifikasi.

Penelitian yang dilakukan oleh (Devella, Yohannes, and Rahmawati 2020) yang berjudul “Implementasi *Random Forest* Untuk Klasifikasi Motif Songket Palembang Berdasarkan SIFT” permasalahan yang dihadapi adalah untuk membedakan motif songket Palembang yaitu dengan cara konvensional dengan memerhatikan bentuk, warna, prinsip dan cara pembuatan. Sehingga peneliti ini menggunakan metode *random forest* berdasarkan sift untuk melakukan klasifikasi motif songket Palembang. Hasil evaluasi model yang dilakukan di penelitian ini meliputi. Motif Bunga Cina nilai *class accuracy* sebesar 91,11%, *precision* 100%, dan *recall* sebesar 69,23%. Selanjutnya motif songket cantik manis dengan *accuracy* sebesar 100%, *precision* 100%, dan *recall* sebesar 100%, dan terakhir motif songket Pulir dengan nilai *accuracy* sebesar 91,11%, *precision* sebesar 78,94%, dan *recall* sebesar 100%.

Penelitian yang dilakukan oleh (Yusuf et al. 2020) yang berjudul “Implementasi Algoritma Naive Bayes dan *Random Forest* Dalam Memprediksi Prestasi Akademik Mahasiswa Universitas Islam Negeri Ar-Raniry Banda Aceh” permasalahan yang dihadapi adalah Prestasi akademik ditentukan oleh dua faktor, yaitu faktor internal yang berasal dari dalam diri individu dalam hal ini mahasiswa dan faktor eksternal yang berasal dari luar diri individu atau hal yang dipengaruhi oleh lingkungan. Ada banyak cara mencari suatu prestasi akademik, salah satunya menggunakan data mining yang bertujuan memprediksi atau mengklasifikasikan data menggunakan algoritma klasifikasi. Sehingga peneliti menggunakan metode *naive bayes* dan *random forest* yang digunakan untuk mengklasifikasikan atau memprediksi prestasi akademik

mahasiswa pada Universitas Islam Negeri Ar-Raniry Banda Aceh. Dari grafik diatas dapat disimpulkan bahwa metode Naive Bayes lebih unggul dikarenakan nilai keakuratan Naive Bayes lebih besar dibandingkan metode *Random Forest*. Nilai tertinggi Naive Bayes sebesar 78,0% sedangkan *Random Forest* sebesar 77,2%.

Penelitian yang dilakukan oleh (Bawono and Wasono 2019) yang berjudul “Perbandingan Metode *Random Forest* Dan Naïve Bayes Untuk Klasifikasi Debitur Berdasarkan Kualitas Kredit” permasalahan yang dihadapi adalah tidak semua kredit dapat dikembalikan secara sempurna dan tepat waktu artinya akan muncul suatu resiko yang dikenal dengan resiko kredit dimana risiko kredit dapat terjadi pada setiap bank. Sehingga peneliti menggunakan metode naïve bayes dan *random forest* untuk mengklasifikasikan debitur berdasarkan kualitas kredit. Hasil yang didapatkan dari pengujian klasifikasi ini antara lain *Random Forest* memiliki hasil akurasi tertinggi yakni sebesar 98,16% disusul Naïve Bayes 95,93%.

Berdasarkan 5 penelitian relevan diatas dapat diambil kesimpulan bahwa penelitian terdahulu membahas tentang metode yang sejenis yaitu dengan menggunakan *random forest*. Kelebihan dari kelima penelitian sebelumnya adalah memudahkan dalam menyelesaikan berbagai masalah dengan menggunakan teknik pada *data mining* yaitu klasifikasi. Kekurangan penelitian yang terdahulu adalah hasil yang didapatkan berupa model dan belum di implemantasikan kedalam *website*.

Pada penelitian ini, memfokuskan pada metode *random forest* untuk proses klasifikasi. Penelitian ini diharapkan dapat bermanfaat pada pemerintah desa dalam proses seleksi penerima bantuan keluarga harapan adapun dibuat sistem sebagai *tools* yang dapat mempermudah dalam penggunaan model *mechine learning* yang dibuat dengan metode *random forest* tersebut.

2.2 Landasan Teori

2.2.1 Kemiskinan

Definisi kemiskinan yang tertuang dalam Peraturan Presiden Republik Indonesia Nomor 7 Tahun 2005 tentang Rencana Pembangunan Jangka Menengah Nasional (RPJMN) 2004-2009 bahwa kemiskinan merupakan suatu kondisi dimana seseorang atau sekelompok orang, laki-laki dan perempuan, tidak terpenuhi hak-hak dasarnya untuk mempertahankan dan mengembangkan kehidupan yang bermartabat. Sedangkan definisi kemiskinan menurut Badan Pusat Statistik (BPS) dan Departemen Sosial, kemiskinan adalah ketidakmampuan individu dalam memenuhi kebutuhan dasar minimal untuk hidup layak (baik makanan maupun non makanan) (Nuzula, Prahutama, and Hakim 2020).

Kemiskinan merupakan salah satu masalah yang dialami oleh beberapa negara berkembang, termasuk Indonesia. Banyak cara yang dilakukan untuk menanggulangi kemiskinan, diantaranya dengan program bantuan sosial untuk rakyat miskin (Ermawati 2019). Berdasarkan penjelasan tersebut bisa disimpulkan bahwa kemiskinan adalah masalah yang dihadapi banyak negara berkembang dan menjadi pusat perhatian di negara manapun.

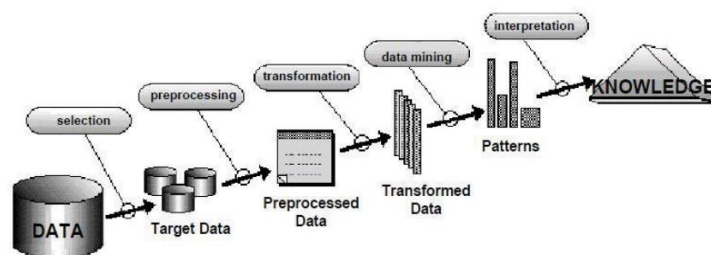
2.2.2 Program Keluarga Harapan

PKH adalah salah satu program bantuan dari pemerintah yang melibatkan semua anggota keluarga, kepala rumah tangga harus memenuhi kualifikasi tertentu yang akan dipilih untuk mendapatkan bantuan tersebut, kualifikasi atau syarat untuk rumah tangga yang bisa mendapatkan bantuan tersebut dapat dibandingkan dengan melihat luas tanah sempit, jenis lantai bangunan, jenis dinding bangunan, wc, sumber air yang digunakan untuk sehari-hari, sumber penerangan dirumah, bahan bakar yang digunakan untuk memasak, mampu atau tidak membayar anggota keluarga ke puskesmas, memiliki atau tidak ibu hamil atau anak balita, dan apakah memiliki anak yang sedang bersekolah. Dari beberapa kriteria tersebut rumah tangga yang berhak atau memenuhi, akan mendapat bantuan sesuai dengan anggaran

biaya pemerintah pertahunnya. Dana tersebut akan diberikan 4kali dalam setahun, para penerima bantuan PKH ini berhak memenuhi sebuah kewajiban setiap bulannya bahwa dana tersebut digunakan untuk semestinya (Pamungkas, Muflikhah, and Wihandika 2019).

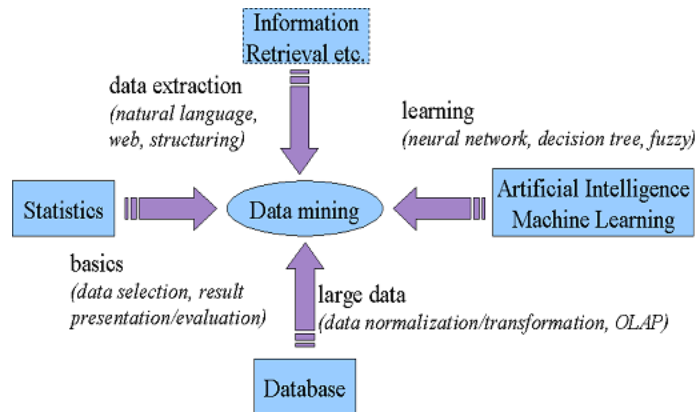
2.2.3 Data Mining

Data mining merupakan teknik untuk mendapatkan manfaat lebih dari data yang sudah ada. *Data mining* bisa diartikan serangkaian proses mendapatkan pengetahuan atau pola dari kumpulan data. Tujuan data mining adalah untuk melakukan klasifikasi, klusterisasi, menemukan pola asosiasi hingga melakukan peramalan (*predicting*) (Ermawati 2019). Istilah *data mining* lebih populer dengan sebutan KDD (*Knowledge Discovery from Database*). Proses KDD dapat dilihat pada gambar 2.1 mulai dari pemilihan atribut data hingga terciptalah sebuah pengetahuan (Antara et al. 2015).



Gambar 2. 1 KDD Process(Sumber : (Antara et al. 2015))

Data mining bukanlah suatu bidang yang baru. Salah satu kesulitan untuk mendefinisikan *data mining* adalah kenyataan bahwa *data mining* mewarisi banyak aspek dan teknik dari bidang-bidang ilmu yang dulu sudah mapan terlebih dulu. Gambar 2.2 menunjukkan bahwa data mining memiliki akar yang panjang dari bidang ilmu yang berbeda seperti kecerdasan buatan *artificial intelligent*, *machine learning*, statistik, *database*, dan juga *information retrieval* (Mardi 2017).



Gambar 2. 2 Bidang ilmu data mining (Sumber (Mardi 2017))

2.2.4 Pengelompokan Data Mining

Menurut (Mustafa, Ramadhan, and Thenata 2013) Ada beberapa teknik yang dimiliki data mining berdasarkan tugas yang bisa dilakukan.

4.1 Deskripsi

Para peneliti biasanya mencoba menemukan cara untuk mendeskripsikan pola dan *trend* yang tersembunyi dalam data.

4.2 Estimasi

Estimasi mirip dengan klasifikasi, kecuali variabel tujuan yang lebih ke arah numerik dari pada kategori.

4.3 Prediksi

Prediksi memiliki kemiripan dengan estimasi dan klasifikasi. Hanya saja, prediksi hasilnya menunjukkan sesuatu yang belum terjadi (mungkin terjadi di masa depan).

4.4 Klasifikasi

Dalam klasifikasi variabel, tujuan bersifat kategorik. Misalnya, kita akan mengklasifikasikan pendapatan dalam tiga kelas, yaitu pendapatan tinggi, pendapatan sedang, dan pendapatan rendah.

4.5 Clustering

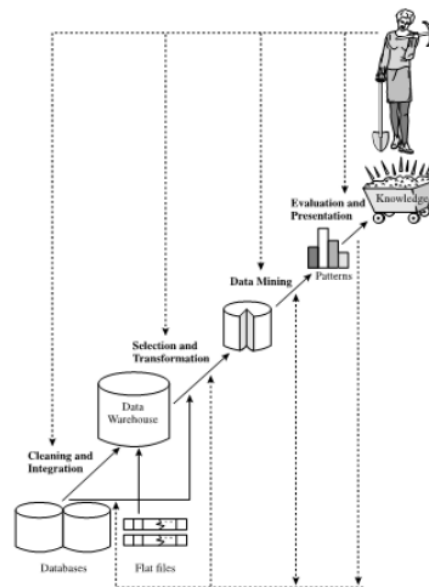
Clustering lebih ke arah pengelompokan *record*, pengamatan, atau kasus dalam kelas yang memiliki kemiripan.

4.6 Asosiasi

Mengidentifikasi hubungan antara berbagai peristiwa yang terjadi pada satu waktu.

2.2.5 Tahap Tahab Data Mining

Tahapan yang dilakukan pada proses data mining diawali dari seleksi data dari data sumber ke data target, tahap *pre-processing* untuk memperbaiki kualitas data, *transformasi*, *data mining* serta tahap interpretasi dan evaluasi yang menghasilkan *output* berupa pengetahuan baru yang diharapkan memberikan kontribusi yang lebih baik (Fatmawati and Windarto 2018). Dijelaskan secara detail pada gambar 2.3.



Gambar 2. 3 Tahab tahab data mining (Sumber (Fatmawati and Windarto 2018))

Berikut penjelasan tahapan data mining berdasarkan gambar 2.3 menurut (Fatmawati and Windarto 2018).

5.1 *Data selection*

Pemilihan (seleksi) data dari sekumpulan data operasional perlu dilakukan sebelum tahap penggalian informasi dalam KDD dimulai. Data hasil seleksi yang digunakan untuk proses *data mining*, disimpan dalam suatu berkas, terpisah dari basis data operasional.

5.2 *Pre-processing /cleaning*

Sebelum proses *data mining* dapat dilaksanakan, perlu dilakukan proses *cleaning* pada data yang menjadi fokus KDD. Proses *cleaning* mencakup antara lain membuang duplikasi data, memeriksa data yang inkonsisten, dan memperbaiki kesalahan pada data.

5.3 *Transformation*

Coding adalah proses *transformasi* pada data yang dipilih, sehingga data tersebut sesuai untuk proses *data mining*. Proses coding dalam KDD merupakan proses kreatif dan sangat tergantung pada jenis atau pola informasi yang akan dicari dalam basis data.

5.4 *Data mining*

Data mining adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Teknik, metode, algoritma dalam *data mining* sangat bervariasi. Pemilihan metode atau algoritma yang tepat sangat bergantung pada tujuan dan proses KDD secara keseluruhan.

5.5 *Interpretation/evaluation*

Pola informasi yang dihasilkan dari proses data mining perlu ditampilkan dalam bentuk yang mudah dimengerti oleh pihak yang berkepentingan. Tahap ini merupakan bagian dari proses KDD yang disebut *interpretation*. Tahap ini mencakup pemeriksaan apakah pola atau informasi yang ditemukan bertentangan dengan fakta atau hipotesis yang ada sebelumnya.

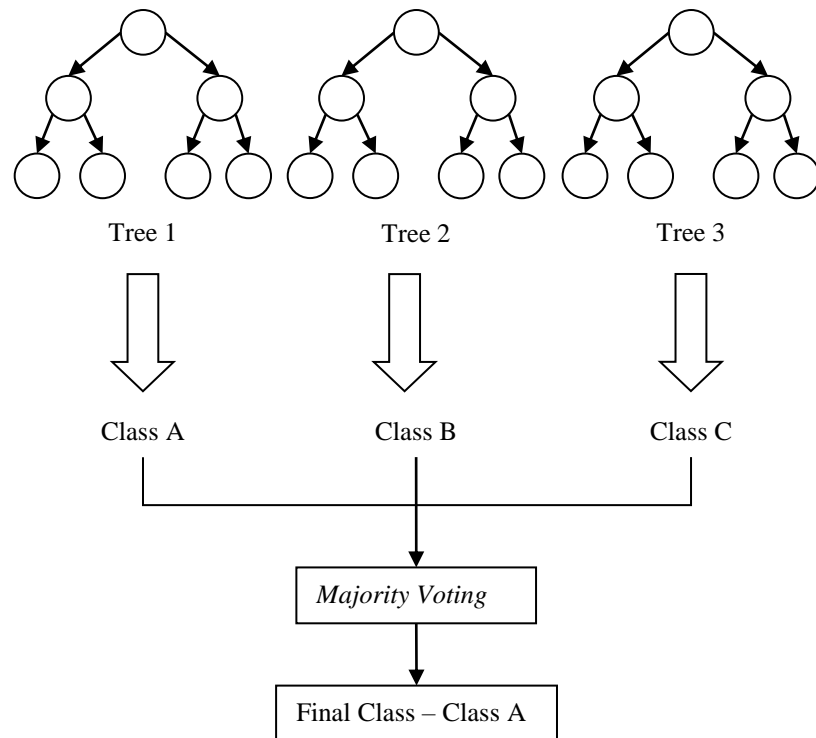
2.2.6 Klasifikasi

Klasifikasi dapat digambarkan sebagai berikut. *Data input*, disebut juga *training set*, terdiri atas banyak contoh (*record*), yang masing-masing memiliki beberapa atribut. Selanjutnya, tiap contoh diberi sebuah label *class* khusus. Tujuannya untuk menganalisa *data input* dan mengembangkan

deskripsi atau model akurat untuk tiap *class* menggunakan fitur-fitur pada data. Deskripsi *class* ini digunakan untuk mengklasifikasikan data pengujian lainnya dengan label *class* tidak diketahui. Deskripsi tersebut juga dapat digunakan untuk memahami tiap *class* dalam data (Rizki et al. 2020).

2.2.7 *Random Forest*

Random Forest merupakan salah satu metode CART (*Classification and Regression Tree*) dalam data mining dan tidak memerlukan asumsi apapun. Metode ini menggunakan konsep pohon keputusan (*decision tree*). Model ini dibentuk dari banyak pohon sehingga membentuk sebuah kumpulan pohon seperti hutan (*forest*) dengan menerapkan metode *bootstrap aggregating (bagging)* dan *random feature selection* (Iman and Wahyu 2021). Ilustrasi dari pohon keputusan dan pengambilan keputusan dalam metode *random forest* dapat dilihat dalam gambar 2.4.



Gambar 2. 4 Ilustrasi metode *random forest* sumber (Iman and Wahyu 2021)

Menurut (Wuisan 2020) Berikut tahapan algoritma dalam pembuatan pohon keputusan dengan Algoritma *Random Forest*:

1. Perhatikan label pada data, jika sudah sama semua, maka akan dibentuk daun dengan nilai label data keseluruhan.
2. Menghitung nilai informasi dengan menggunakan semua data yang ada dengan formula:

$$Info(D) = - \sum_{i=1}^m p_i \log^2(p_i) \dots\dots\dots (2.1)$$

Dimana merupakan probabilitas tuple dalam D yang menjadi kelas dengan asumsi atau disebut juga *entropy* dari D merupakan rata rata informasi yang diperlukan untuk identifikasi tuple dalam D.

Jika Nilai A adalah nilai diskrit maka data D akan dipisahkan sejumlah nilai data A sehingga nilai setiap cabang akan murni dan sejenis. Setelah percabangan pertama, jumlah percabangan yang mungkin terjadi diukur dengan persamaan:

$$infoA(D) \sum_j^v \frac{|D_j|}{|D|} x InfoA(D_j) \dots\dots\dots (2.2)$$

3. Menghitung nilai informasi dengan formula
4. Untuk setiap atribut dengan memperhatikan isi data dari atribut. Dimana $\frac{|D_j|}{|D|}$ merupakan bobot dari partisi j. *infoA(D)* merupakan informasi yang diperlukan untuk mengklasifikasi *tuple* dari D pada partisi A. Semakin kecil hasil persamaan ini, semakin baik pula partisi yang dihasilkan. Nilai dari sebuah atribut menentukan penting tidaknya atribut tersebut dalam penyusunan pohon keputusan. Jika atribut bernilai kontinyu, maka akan dicari *split_point* dengan cara mengurutkan seluruh data menurut atribut tersebut dari kecil ke besar, lalu di rata-rata antar satu data dengan data setelahnya. Nilai informasi akan dihitung menurut satu persatu calon *split_point* dan nilai *split_point* yang akan dipilih yang terkecil. Nilai *gain* untuk

setiap atribut akan diperhitungkan dengan formula (2.3), nilai dengan gain tertinggi akan dijadikan cabang dalam pohon keputusan.

$$GAIN(A) = Info(D) - InfoA(D) \dots\dots\dots (2.3)$$

5. Setelah cabang pohon keputusan terbentuk, perhitungan dilakukan kembali seperti pada tahap 1 sampai 4. Namun jika cabang telah mencapai maksimal cabang yang diperbolehkan, daun akan terbentuk dengan nilai mayoritas dari nilai data.

2.2.8 Python

Pengertian Python (bahasa pemrograman) merupakan bahasa pemrograman tinggi yang bisa melakukan eksekusi sejumlah instruksi multi guna secara langsung (interpretatif) dengan metode *Object Oriented Programming* dan juga menggunakan semantik dinamis untuk memberikan tingkat keterbacaan *syntax*. Sebagai bahasa pemrograman tinggi, python dapat dipelajari dengan mudah karena telah dilengkapi dengan manajemen memori otomatis (Rahmadhika and Thantawi 2021).

2.2.9 Flask

Flask adalah sebuah *web framework* yang ditulis dengan bahasa Python dan tergolong sebagai jenis *microframework*. Flask berfungsi sebagai kerangka kerja aplikasi dan tampilan dari suatu *web*. Dengan menggunakan Flask dan bahasa Python, pengembang dapat membuat sebuah *web* yang terstruktur dan dapat mengatur behaviour suatu *web* dengan lebih mudah (Irsyad 2018).

2.2.10 K-fold Cross Validation

K-fold *Cross Validation* adalah salah satu cara menemukan parameter terbaik dari satu model dengan cara menguji besarnya error pada data test. Dalam *cross validation* data dibagi menjadi k - subset dengan ukuran yang sama. Kemudian subset tersebut akan dibagi menjadi data train dan data test. Tujuan dari *cross validation* adalah untuk mendefinisikan

dataset untuk "menguji" model dalam tahap pelatihan (yaitu, validasi data), dalam rangka untuk membatasi masalah seperti terjadinya *overfitting*, memberikan wawasan tentang bagaimana model akan menggeneralisasi independen dataset dan lain-lain(Indah Prabawati, Widodo, and Ajie 2019).

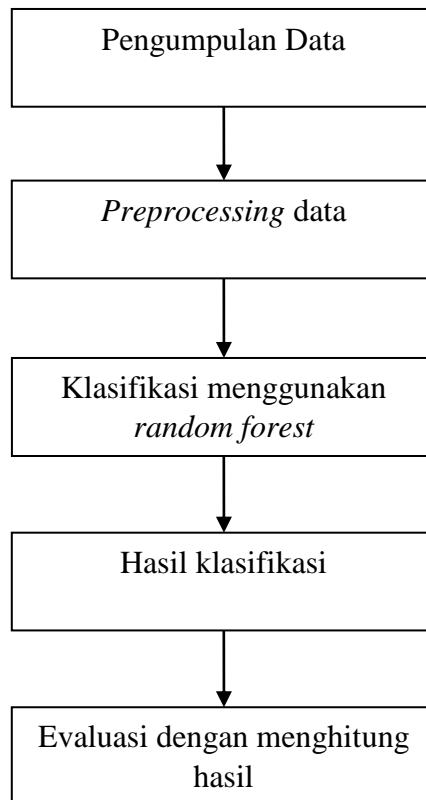
2.2.11 *Synthetic Minority Over-sampling Technique*

Synthetic Minority Over-sampling Technique (SMOTE) merupakan metode yang populer diterapkan dalam rangka menangani ketidak seimbangan kelas. Teknik ini mensintesis sampel baru dari kelas minoritas untuk menyeimbangkan dataset dengan cara sampling ulang sampel kelas minoritas (Siringoringo 2018).

BAB 3 METODOLOGI PENELITIAN

3.1 Prosedur Penelitian

Prosedur yang dilakukan dalam penelitian ini dimulai dengan menentukan alur kerja *random forest* untuk mendapatkan hasil yang akurat. Alur penelitian bisa dilihat pada gambar 3.1.



Gambar 3. 1 Kerangka kerja penelitian

Uraian prosedur penelitian yang digambarkan pada gambar 3.1 adalah sebagai berikut:

1. Pengumpulan Data

Pada langkah awal penelitian ini data yang digunakan merupakan data dari Kantor Kelurahan Desa Balerejo Kecamatan Kaliangakrik Kabupaten Magelang. Dari data tersebut memiliki atribut nama_KRT, status_bangunan, status_lahan, luas_lantai, lantai, dinding, kondisi dinding,

atap, kondisi_atap, sumber_air_minum, cara_memperoleh_airminum, ibu_hamil, anak_sekolah, balita, lansia, penerima PKH.

2. Preprocessing Data

Preprocessing dilakukan untuk membersihkan data yang bertujuan agar data yang digunakan telah bebas dari *noise* sehingga data yang didapat kualitasnya baik. Dalam penelitian ini akan dilakukan partisi data yang digunakan sebagai perbandingan menggunakan *percentage split* dan *cross validation* dengan jumlah yang berdeda beda *percentase spilt* menggunakan 20%, 30%, 40%, 50% dan *cross validaton* menggunakan 5-fold, 7-fold, 10-fold, 12-fold.

3. Klasifikasi Menggunakan Random Forest

Tahapan klasifikasi dilakukan dengan mengklasifikasian data yang telah didapat dari Desa dengan menggunakan metode *random forest*, yang bertujuan agar mendapatkan *role* pohon keputusan yang nantinya dipakai untuk membuat sistem pengelolaan kelayakan penerima bantuan PKH.

4. Hasil Klasifikasi

Setelah diklasifikasikan hasil akan ditampilkan di halaman hasil yang akan dibuat didalam sistem pengelolaan kelayakan penerima bantuan PKH.

5. Evaluasi Dengan Menghitung Hasil

Setelah dilakukan pengujian dan mendapatkan hasil langkah selanjutnya adalah evaluasi. Pada tahap ini akan dihitung akurasi terhadap pengujian klasifikasi menggunakan *random forest* dengan menggunakan *confusion matrix* dengan perhitungan seperti gambar 3.2.

		<i>Predicted</i>	
		<i>Negative</i>	<i>Positive</i>
<i>Actual</i>	<i>Negative</i>	TN	FP
	<i>Positive</i>	FN	TP

Gambar 3. 2 *confusion matrix* (Sumber (Setyawan and Suradi 2017))

Keterangan:

TN (*True Negatif*): jumlah data testing yang sebenarnya negatif dan terprediksi negatif.

TP (*True Positif*): jumlah data testing yang sebenarnya positif dan terprediksi positif.

FN (*False Negatif*): jumlah data testing yang sebenarnya positif dan terprediksi negatif.

FP (*False Positif*): jumlah data testing sebenarnya negatif dan terprediksi positif.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \dots\dots\dots (3.1)$$

$$Precision = \frac{TP}{TP + FP} \dots\dots\dots (3.2)$$

$$Recall = \frac{TP}{FN + TP} \dots\dots\dots (3.3)$$

Dari hasil penelitian yang dilakukan ini nanti akan menghasilkan sebuah klasifikasi untuk penerima bantuan PKH agar lebih tepat sasaran, implementasi setelah mendapatkan model random forest akan dibuat sistem agar model tersebut mudah dalam penggunaannya berikut rancang *interface* yang akan digunakan:

Halaman *Login*

The image shows a simple login form within a rectangular border. At the top center, the text "Halaman Login" is displayed. Below it, the label "Username" is positioned above a horizontal input field. Further down, the label "Password" is positioned above another horizontal input field. At the bottom center, there is a rounded rectangular button labeled "Login".

Gambar 3. 3 Halaman Login

Gambar 3.3 Merupakan Desain untuk *login* pegawai yang digunakan untuk *login* kedalam sistem

Halaman *Home*

The image shows a home page layout. At the top, there is a navigation bar with the following items: "Logo", "Home", "Data", "Klasifikasi", "About", and "Log Out". Below the navigation bar, the text "Selamat Datang" is centered. Underneath, there are three data boxes arranged horizontally. Each box has a header section and a data section. The first box is titled "Jumlah Penduduk Miskin", the second is "Jumlah Penerima PKH", and the third is "Jumlah Bukan Penerima PKH". Each box has a horizontal line separating the header from the data area, which is currently empty.

Gambar 3. 4 Halaman *Home*

Gambar 3.4 terdapat desain halaman *home* atau halaman utama dihalaman ini pegawai desa dapat melihat jumlah penduduk miskin, penerima PKH dan bukan penerima PKH.

Halaman Data

Logo	Home	Data	Klasifikasi	About	Log Out
Data Penduduk Miskin					
<input type="button" value="Tambah Data"/>					
no	nama				Aksi
					<input type="button" value="edit"/> <input type="button" value="Hapus"/>

Gambar 3. 5 Halaman Data

Gambar 3.5 terdapat halaman data di halaman tersebut pegawai desa dapat menambahkan data, menghapus data dan mengedit data.

Halaman Klasifikasi

Logo	Home	Data	Klasifikasi	About	Log Out
Kelayakan Penerima PKH					
Status Bangunan					
<input type="text"/>					
Status lahan					
<input type="text"/>					
<input type="button" value="Klasifikasi"/>					

Gambar 3. 6 Halaman Klasifikasi

Gambar 3.6 terdapat halaman proses dimana pegawai desa dapat mengklasifikasikan kelayakan penerima bantuan PKH dengan cara mengisi atribut yang ada.



Gambar 3. 7 Halaman Hasil

Gambar 3.7 adalah halaman hasil, halaman tersebut menampilkan hasil dari proses klasifikasi yang ada pada pada halaman klasifikasi dengan menginputkan nama dan kriteria keadaan penduduk.

3.2 Metode Pengumpulan Data

1. Observasi

Pada penelitian ini metode pengumpulan data dengan meminta data resmi pada Kantor Kelurahan Desa Balerejo, Kecamatan Balerejo, Kabupaten Magelang. Data yang didapat dalam kelurahan tersebut sebanyak 504 data. Dari data tersebut terdapat beberapa atribut antara lain nama_KRT, status_bangunan, status_lahan, luas_lantai, lantai, dinding, kondisi dinding, atap, kondisi_atap, sumber_air_minum, cara_memperoleh_airminum, ibu_hamil, anak_sekolah, balita, lansia, penerima PKH.

2. Wawancara

Wawancara dilakukan dengan kepala desa dan perangkat desa adapaun hasil yang didapatkan saat wawancara:

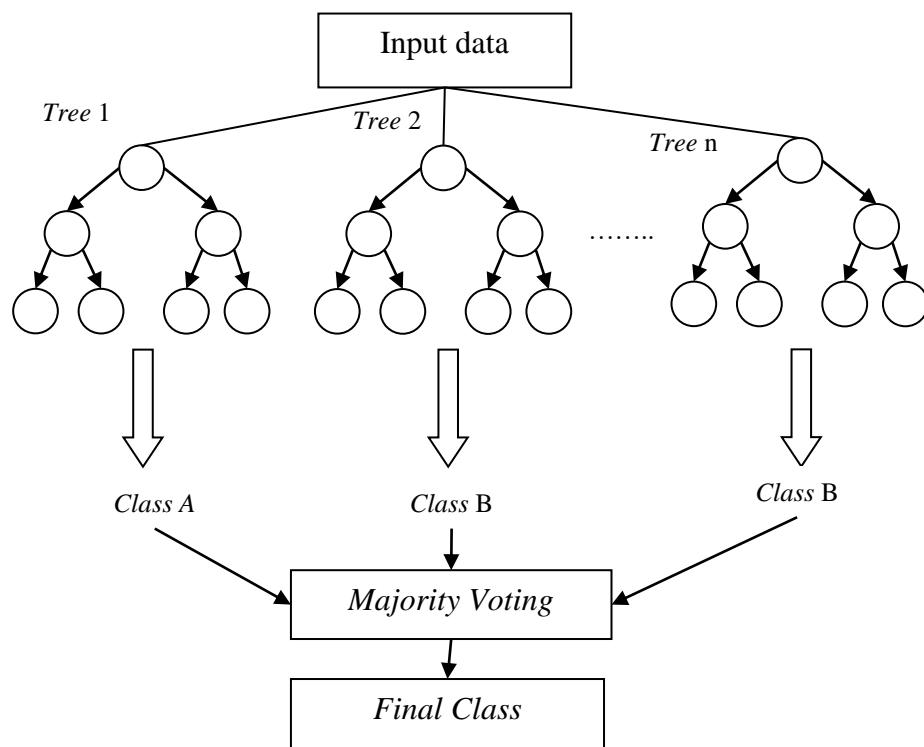
- a. Kriteria yaitu faktor penentu penerima bantuan keluarga harapan yang selama ini dilakukan di Desa Balerejo. Kriteria yang digunakan adalah komponen penerima PKH antara lain ibu hamil, anak sekolah, balita dan lansia.

b. Proses pemberian bantuan PKH di Desa Balerejo

Berdasarkan informasi yang didapat proses pemberian bantuan di Desa Balerejo adalah data yang didapat dari pusat kemudian diseleksi mempunyai komponen atau tidak, jika mempunyai komponen lanjut jika tidak, tidak mendapatkan. Kemudian jika merasa layak tapi tidak masuk daftar penerima bantuan PKH bisa mengajukan dikantor kelurahan Desa Balerejo.

3.3 Metode Pengolahan Data

Untuk pengolahan pada penelitian ini menggunakan *random forest*. Metode pengolahan data yang dilakukan dapat dilihat pada gambar 3.8.



Gambar 3. 8 Konsep sederhana *random forest* (Sumber(Danoedoro and Murti 2021))

Ketika menginput data, *random forest* akan membuat *subset* dari data tersebut sebanyak iterasi yang dilakukan. *Subset* pertama akan membuat *tree* sendiri (*tree* pertama), *subset* kedua membuat *tree* kedua, dan seterusnya berdasarkan iterasi yang dilakukan. *Tree* tersebut akan menghasilkan kelas-kelas

tersendiri (*class A*, *class B*) dan akan di-voting menggunakan *Majority-Voting* sehingga menghasilkan *Final-Class* (Danoedoro and Murti 2021).

BAB 5

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Dari hasil penelitian klasifikasi kelayakan penerima bantuan PKH di Desa Balerejo Kecamatan Kaliangkrik Kabupaten Magelang dapat disimpulkan sebagai berikut:

1. *Random forest* yang paling tinggi akurasinya yang digunakan untuk model klasifikasi penerima PKH adalah *random forest* yang memiliki 5 pohon.
2. Dalam penentuan penerima bantuan PKH menggunakan *random forest*. Tingkat akurasi tertinggi terdapat pada pengujian dengan *persentase* split yaitu 30% dengan akurasi yang didapat adalah 74.1%.
3. Dengan menggunakan sistem klasifikasi ini diharapkan mempermudah pegawai Desa dalam memberikan bantuan PKH kepada masyarakat yang benar benar layak dengan menginputkan kondisi rumah dan atribut lainnya.

5.2 Saran

Saran yang dapat diberikan untuk mengembangkan sistem pada penelitian ini adalah:

1. Klasifikasi penerima bantuan PKH dapat dilakukan dengan menggunakan metode lain yaitu algoritma C4.5, naïve bayes, KNN, *neural network*.
2. Permasalahan penentuan bantuan PKH dapat dikembangkan menggunakan aplikasi berbasis *mobile* agar lebih mudah digunakan karena bisa langsung diakses di *smartphone* pegawai.

DAFTAR PUSTAKA

- Anggraeni, Nurul, Ghina Arifiana, and Agus Maman Abadi. 2017. "Klasifikasi Kualitas Air Sungai Winongo Menggunakan Fuzzy Inference System (FIS) Metode Mamdani." *Program Studi Matematika FMIPA UNY* (2014): 161–70.
- Antara, Kemiskinan et al. 2015. "PERBANDINGAN AKURASI KLASIFIKASI TINGKAT." 2(1): 37–43.
- Bawono, Bonggo, and Rochdi Wasono. 2019. "Perbandingan Metode Random Forest Dan Naive Bayes." *Jurnal Sains dan Sistem Informasi* 3(7): 343–48.
- Danoedoro, Projo, and Sigit Heru Murti. 2021. "Klasifikasi Tutupan Lahan Data Landsat-8 Oli Menggunakan Metode Random Forest." *Jurnal Penginderaan Jauh Indonesia* 03(01): 1–7.
- Devella, Siska, Yohannes Yohannes, and Firda Novia Rahmawati. 2020. "Implementasi Random Forest Untuk Klasifikasi Motif Songket Palembang Berdasarkan SIFT." *JATISI (Jurnal Teknik Informatika dan Sistem Informasi)* 7(2): 310–20.
- Ermawati, Erni. 2019. "Algoritma Klasifikasi C4.5 Berbasis Particle Swarm Optimization Untuk Prediksi Penerima Bantuan Pangan Non Tunai." *Sistemasi* 8(3): 513.
- Fatmawati, Kiki, and Agus Perdana Windarto. 2018. "Data Mining: Penerapan Rapidminer Dengan K-Means Cluster Pada Daerah Terjangkit Demam Berdarah Dengue (Dbd) Berdasarkan Provinsi." *Computer Engineering, Science and System Journal* 3(2): 173.
- Iman, Qonita, and Arie Wahyu. 2021. "Klasifikasi Rumah Tangga Penerima Beras Miskin (Raskin)/ Beras Sejahtera (Rastra) Di Provinsi Jawa Barat Tahun 2017 Dengan Metode Random Forest Dan Support Vector Machine Classification of Poor Rice (Raskin)/ Prosperous Rice (Rastra) Recipient Hou." 09(2): 178–84.
- Indah Prabawati, Nurul, Widodo, and Hamidillah Ajie. 2019. "Kinerja Algoritma Classification And Regression Tree (Cart) Dalam Mengklasifikasikan Lama Masa Studi Mahasiswa Yang Mengikuti Organisasi Di Universitas Negeri Jakarta." *PINTER : Jurnal Pendidikan Teknik Informatika dan Komputer* 3(2): 139–45.
- Irsyad, Rahadian. 2018. "Penggunaan Python Web Framework Flask Untuk Pemula."
- Mardi, Yuli. 2017. "Data Mining : Klasifikasi Menggunakan Algoritma C4.5." *Jurnal Edik Informatika* 2(2): 213–19.

- Mustafa, M. Syukri, Muh Rizky Ramadhan, and Angelina P. Thenata. 2013. "Implementasi Data Mining Untuk Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naive Bayes Classifier." *Creative Information Technology Journal* 4(2): 151.
- Nuzula, Lutfia, Alan Prahutama, and Arief Rachman Hakim. 2020. "KLASIFIKASI STATUS KEMISKINAN RUMAH TANGGA DENGAN METODE SUPPORT VECTOR MACHINES (SVM) DAN CLASSIFICATION AND REGRESSION TREES (CART) MENGGUNAKAN GUI R (Studi Kasus Di Kabupaten Wonosobo Tahun 2018)." *Jurnal Gaussian* 9(4): 525–34.
- Pamungkas, Vidya Capristyan, Lailil Muflikhah, and Randy Cahya Wihandika. 2019. "Klasifikasi Penerimaan Program Keluarga Harapan (PKH) Menggunakan Metode Learning Vector Quantization (Studi Kasus Desa Kedungjati)." *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer* 3(3): 2659–66.
- Rahmadhika, Muhammad Karunia, and Ahmad Muhammad Thantawi. 2021. "Rancang Bangun Aplikasi Face Recognition Pada Pendekatan CRM Menggunakan Opencv Dan Algoritma Haarcascade." *IKRA-ITH INFORMATIKA: Jurnal Komputer dan Informatika* 5(1): 109–18.
- Rizki, Mhd Yuda et al. 2020. "Implementasi C45 Dalam Memprediksi Index Prestasi Mahasiswa / i Menurut Kebiasaan Belajar." *Seminar Nasional Teknologi Komputer & Sains (SAINTEKS)*: 297–303.
- Sataloff, Robert T, Michael M Johns, and Karen M Kost. *Kecamatan Kaliangkrik Dalam Angka*.
- Setyawan, Doni, and Agustinus Suradi. 2017. "Implementasi Web Service Dan Analisis Kinerja Algoritma Klasifikasi Data Mining Untuk Memprediksi Diabetes Mellitus." *Simetris: Jurnal Teknik Mesin, Elektro dan Ilmu Komputer* 8(2): 701.
- Siringoringo, Rimbun. 2018. "Klasifikasi Data Tidak Seimbang Menggunakan Algoritma SMOTE Dan K-Nearest Neighbor." *Jurnal ISD* 3(1): 44–49.
- Sugianto, Castaka Agus, and Firdi Rizky Maulana. 2019. "Algoritma Naïve Bayes Untuk Klasifikasi Penerima Bantuan Pangan Non Tunai (Studi Kasus Kelurahan Utama)." *Techno.Com* 18(4): 321–31.
- Wuisan, Dewi S.S. 2020. "Journal of Technology Information." *Journal Of Technology Information* 6(1): 29–34.
- Yusuf, Bustami et al. 2020. "Implementasi Algoritma Naive Bayes Dan Random Forest Dalam Memprediksi Prestasi Akademik Mahasiswa Universitas Islam Negeri Ar-Raniry Banda Aceh." *Cyberspace: Jurnal Pendidikan Teknologi Informasi* 4(1): 50.
- Yusuf Sulisty Nugroho, Nova Emiliyawati. 2017. "Sistem Klasifikasi Variabel

Tingkat Penerimaan Konsumen Terhadap Mobil Menggunakan Metode Random Forest.” *Jurnal Teknik Elektro* 9(1): 24–29.