



Available online at : <http://bit.ly/InfoTekJar>

InfoTekJar : Jurnal Nasional Informatika dan Teknologi Jaringan

ISSN (Print) 2540-7597 | ISSN (Online) 2540-7600



Komparasi Dua Metode Algoritma Klasifikasi Untuk Prediksi Pemberian Kartu Jakarta Pintar

Agustienna Merdekawati, Jefina Tri Kumalasari

Program Studi Sistem Informasi, Universitas Bina Sarana Informatika, Jakarta

KEYWORDS

Pemberian KJP, Prediksi, Algoritma C4.5, ID3, Algoritma *Naive Bayes*

CORRESPONDENCE

Phone: 083875819011

E-mail: agustienna.atd@bsi.ac.id

A B S T R A C T

Mendapatkan program Kartu Jakarta Pintar (KJP) harus memenuhi berbagai persyaratan dan kriteria, sayangnya proses penerimaan ini masih dilakukan secara subjektif sehingga rentan untuk tepat sasaran. Knowledge Discovery in Database (KDD) diperlukan dalam penentuan penerima KJP dengan menemukan pola prediksi terbaik. Penelitian ini membandingkan algoritma klasifikasi yaitu ID3 dan *Naive Bayes* guna mengekstrak data kemudian mencari model yang sesuai dalam penentuan proses penerimaan KJP dengan menggunakan sekelompok data sehingga didapatkan persentase precision, recall dan accuracy. Keduanya memiliki proses awal yang sama yaitu *pre processing* atau *data cleaning* guna membuang data yang tidak sesuai baik data kosong maupun tidak konsisten. Pada algoritma ID3 digunakan pohon keputusan dimana sebelumnya diperlukan pencarian *entropi* dan *gain* sedangkan pada *Naive Bayes* dengan menghitung jumlah class. Hasil Analisa keduanya kemudian dilakukan proses pengujian dilakukan untuk membandingkan tingkat akurasi data dengan menerapkan *confusion matrix* dan visualisasi kurva ROC. Hasil pengujian didapat algoritma ID3 menunjukkan tingkat akurasi yang lebih tinggi 77,78% setelah dibandingkan dengan *Naive Bayes*. Tools yang digunakan dalam pengolahan data ini yaitu Rapid Miner.

PENDAHULUAN

Pendidikan merupakan suatu hajat yang penting untuk mendukung kualitas suatu sumber daya manusia. Suatu negara akan meningkat maju jika kualitas pendidikannya sangat baik. Guna dari pendidikan yaitu dapat mencerdaskan anak bangsa, dengan mencerdaskan anak bangsa maka kesejahteraan rakyat akan tercapai dan siap menghadapi perkembangan zaman. Pemerintah DKI Jakarta telah memberikan himbauan mengenai wajib belajar 12 tahun, namun masih banyak anak yang putus sekolah karena ketidakmampuan kebutuhan biaya pendidikan. Berdasarkan data [1] indikator pendidikan DKI Jakarta tingkat sekolah dasar dan SMA atau SMK mengalami peningkatan dalam anak yang putus sekolah, anak yang mengalami putus sekolah pada sekolah dasar meningkat 0,51% naik menjadi 0,69% dan anak yang putus sekolah pada tingkat SMA/SMK meningkat sebesar 0,76% menjadi 0,84%, berikut grafik peningkatan putus sekolah tingkat SD pada gambar 1. dan grafik peningkatan putus sekolah tingkat SMA atau SMK pada gambar 2:



Gambar 1. Grafik Peningkatan Putus Sekolah Tingkat SD



Gambar 2. Grafik Peningkatan Putus Sekolah Tingkat SMA atau SMK

Dari permasalahan diatas, pemerintah memberikan suatu program subsidi pendidikan dengan memberikan keringanan biaya pendidikan, seperti Bantuan Operasional Pendidikan (BOP), Bantuan Operasional Sekolah (BOS), Kartu Jakarta Pintar (KJP) Plus, dan Kartu Jakarta Mahasiswa Unggulan [2].

Kartu Jakarta Pintar Plus (KJP Plus) merupakan suatu program yang diberikan untuk warga negara DKI Jakarta untuk kalangan masyarakat yang tidak mampu dalam mendapatkan pendidikan dari sekolah dasar sampai tamat SMA atau SMK melalui dana APBD DKI Jakarta. Dana pendidikan ini berupa kartu ATM yang dapat digunakan untuk membayar biaya pendidikan sekolah, seperti buku, alat tulis, seragam sekolah, dan lainnya.

Penerimaan Kartu Jakarta Pintar (KJP) masih terjadi penyimpangan, sistem penerimaan Kartu Jakarta Pintar (KJP) masih bersifat subjektif, sehingga tidak tepat sasaran [3]. Berdasarkan permasalahan tersebut harus dilakukan evaluasi agar penerimaan Kartu Jakarta Pintar (KJP) dapat tepat sasaran dan tercipta keadilan sebagai warga negara. Sehingga warga yang putus sekolah dapat kembali mengenyam pendidikan dengan bantuan pemerintah dan dapat meningkatkan kualitas sumberdaya manusia dikemudiannya. Oleh karena itu, *Naive Bayes* dan Algoritma C4.5 dapat digunakan untuk menyeleksi siswa siswi yang layak mendapatkan Kartu Jakarta Pintar (KJP) sesuai dengan kriteria masing-masing.

Penelitian yang dilakukan [3], metode algoritma *Naive Bayes* menggunakan data *training* agar membentuk peluang kejadian dari setiap kriteria dengan *class* yang berbeda, nilai dari peluang kejadian tersebut dapat digunakan untuk memprediksi suatu keadaan.

Dengan menggunakan algoritma C4.5 memberikan hasil tingkat akurasi cukup tinggi yang didapat dari data *training* dan data *testing* [4].

Dalam penelitian Iskandar dan Suprpto dalam [5], Hasil perbandingan antara algoritma C4.5 dan algoritma *Naive Bayes*, bahwa dengan menggunakan algoritma C4.5 memberikan tingkat akurasi yang lebih tinggi dibandingkan dengan metode algoritma *Naive Bayes*.

METODOLOGI

Pada penelitian ini menggunakan sumber data sekunder. Data sekunder [6] diperoleh dengan lebih cepat dan dengan biaya yang rendah, serta data sekunder diperoleh dari pihak lain atau pihak tempat penelitian, yang sebelumnya data tersebut telah digunakan untuk tujuan tertentu. Namun pada penelitian ini mengambil data pada SMA Negeri 113. Teknik pengambilan data pada penelitian ini menggunakan:

1. Observasi
Melakukan pengamatan langsung bagaimana proses seleksi pemberian Kartu Jakarta Pintar (KJP).
2. Wawancara
Melakukan tanya jawab kepada bagian tata usaha mengenai kriteria yang dijadikan proses seleksi dan permasalahan apa yang terjadi pada saat seleksi sampai kepada permasalahan pemberian Kartu Jakarta Pintar (KJP).
3. Studi Pustaka
Membaca dan mencatat informasi yang dijadikan sumber referensi penelitian, seperti buku, jurnal, artikel, internet, dan lainnya.

Setelah didapatkan data, selanjutnya melakukan pengolahan data mining dengan menggunakan *Knowledge discovery in databases* (KDD). Data mining merupakan proses menemukan pola-pola baru dari sekumpulan data yang sangat besar [7]. Data mining merupakan bagian metode dari *Knowledge Discovery in Databases* (KDD) [8]. Adapun proses KDD sebagai berikut [9]:

- a. *Data Selection*
Data *selection* merupakan proses pemilihan data, dimana penyimpanan data operasional terpisah dengan data mining,

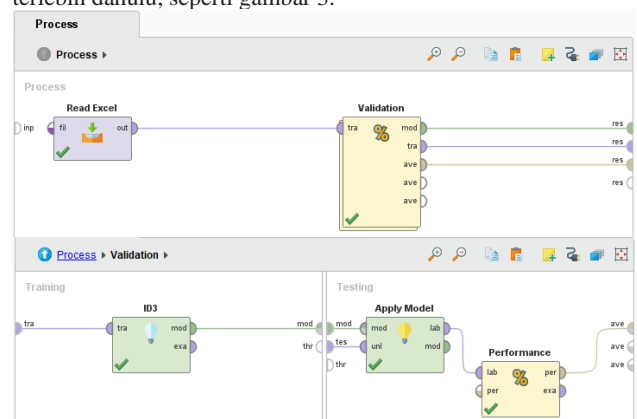
yaitu dengan mengambil kriteria data yang tidak berhubungan dengan penelitian, maka data ini tidak terpakai dan tidak tersimpan dalam data mining. Contoh untuk atribut nama dan alamat.

- b. *Pre-processing/Cleaning*
Proses dari *pre-processing/cleaning* yaitu membuang data yang berganda, memeriksa data yang tidak konsisten, dan memperbaiki kesalahan data. Hasil dari *pre-processing* berjumlah 91 data.
- c. *Transformation*
Pada proses *transformation*, suatu atribut diubah menjadi beberapa atribut. Pada proses ini data telah dipilih dan dapat digunakan dalam proses data mining.
- d. *Data mining*
Setelah melewati tahap *selection*, *pre-processing*, dan *transformation*, selanjutnya dilakukan pengolahan data dengan menggunakan data mining, guna menghasilkan pola-pola atau model prediksi kelayakan pemberian Kartu Jakarta Pintar (KJP), data yang digunakan pada penelitian ini berjumlah 91 data. Pengolahan data mining ini menggunakan metode algoritma C4.5 dan algoritma *naive bayes* dengan menggunakan *Rapidminer*.
- e. *Interpretation/Evaluation*
Pada tahap evaluasi, mengukur tingkat keakuratan hasil dan validasi yang dicapai. Evaluasi tingkat keakuratan ini menggunakan *confusion matrix* dan ROC/AUC (*Area Under Cover*). *Confusion matrix* suatu alat yang digunakan untuk menganalisa tingkat pengklasifikasian dapat mengenali atribut dari class yang berbeda [10]. Hasil dari *Confusion matrix* berupa nilai *accuracy*, *precision*, dan *recall*. Kurva ROC digunakan untuk perbandingan klasifikasi dalam bentuk gambar dan menunjukkan tingkat akurasi [10]. Klasifikasi tingkat akurasi ROC/AUC terdapat lima kelompok, yaitu:
 1. Nilai akurasi 0.90–1.00 menyatakan *Excellent Classification*
 2. Nilai akurasi 0.80–0.90 menyatakan *Good Classification*
 3. Nilai akurasi 0.70 – 0.80 menyatakan *Fair Classification*
 4. Nilai akurasi 0.60 – 0.70 menyatakan *Poor Classification*
 5. Nilai akurasi 0.50 – 0.60 menyatakan *Failure*

HASIL DAN DISKUSI

Pengolahan ID3

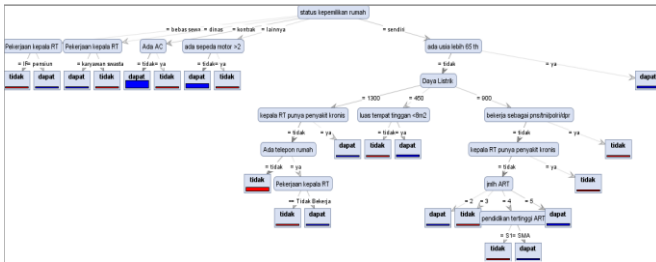
Awal proses menggunakan *rapidminer* dengan membuat desain terlebih dahulu, seperti gambar 3.



Gambar 3. Model Desain Menggunakan Algoritma ID3

Selanjutnya import data excel yang akan diolah ke dalam read excel, lalu *running*-kan. Setelah itu akan ditampilkan rule dan

pohon keputusan. Berikut pohon keputusan dari data pemberian Kartu Jakarta Pintar menggunakan ID3:



Gambar 4. Pohon Keputusan Pemberian KJP Menggunakan ID3

Berikut rule berdasarkan hasil pohon keputusan menggunakan ID3, yaitu:

1. Jika status kepemilikan rumah bebas sewa dan pekerjaan kepala rumah tangga yaitu IRT, maka tidak mendapat KJP
2. Jika status kepemilikan rumah bebas sewa dan pekerjaan kepala rumah tangga yaitu pensiunan, maka mendapatkan KJP
3. Jika status kepemilikan rumah yaitu dinas dan pekerjaan kepala rumah tangga yaitu IRT, maka mendapatkan KJP
4. Jika status kepemilikan rumah yaitu dinas dan pekerjaan kepala rumah tangga yaitu karyawan swasta, maka tidak mendapatkan KJP
5. Jika status kepemilikan rumah yaitu kontrak dan tidak memiliki AC, maka mendapatkan KJP
6. Jika status kepemilikan rumah yaitu kontrak dan memiliki AC, maka tidak mendapatkan KJP
7. Jika status kepemilikan rumah yaitu lainnya dan tidak memiliki sepeda motor >2, maka mendapatkan KJP
8. Jika status kepemilikan rumah yaitu lainnya dan memiliki sepeda motor >2, maka tidak mendapatkan KJP
9. Jika status kepemilikan rumah yaitu milik sendiri, anggota keluarga tidak ada yang usianya >65tahun, daya listrik sebesar 1300, kepala rumah tangga tidak memiliki penyakit kronis, dan tidak memiliki telepon rumah, maka tidak mendapatkan KJP
10. Jika status kepemilikan rumah yaitu milik sendiri, anggota keluarga tidak ada yang usianya >65tahun, daya listrik sebesar 1300, kepala rumah tangga tidak memiliki penyakit kronis, memiliki telepon rumah, dan pekerjaan kepala rumah tangga yaitu IRT, maka tidak mendapatkan KJP
11. Jika status kepemilikan rumah yaitu milik sendiri, anggota keluarga tidak ada yang usianya >65tahun, daya listrik sebesar 1300, kepala rumah tangga tidak memiliki penyakit kronis, memiliki telepon rumah, dan pekerjaan kepala rumah tangga yaitu tidak bekerja, maka mendapatkan KJP
12. Jika status kepemilikan rumah yaitu milik sendiri, anggota keluarga tidak ada yang usianya >65tahun, daya listrik sebesar 1300, kepala rumah tangga memiliki penyakit kronis, maka mendapatkan KJP
13. Jika status kepemilikan rumah yaitu milik sendiri, anggota keluarga tidak ada yang usianya >65tahun, daya listrik sebesar 450, luas tempat tinggal tidak <8m², maka tidak mendapatkan KJP
14. Jika status kepemilikan rumah yaitu milik sendiri, anggota keluarga tidak ada yang usianya >65tahun, daya listrik sebesar 450, luas tempat tinggal <8m², maka mendapatkan KJP
15. Jika status kepemilikan rumah yaitu milik sendiri, anggota keluarga tidak ada yang usianya >65tahun, daya listrik sebesar 900, tidak bekerja sebagai PNS/TNI/Polri/DPR, kepala rumah tangga tidak memiliki penyakit kronis, jumlah anggota keluarga 2, maka mendapatkan KJP
16. Jika status kepemilikan rumah yaitu milik sendiri, anggota keluarga tidak ada yang usianya >65tahun, daya listrik sebesar 900, tidak bekerja sebagai PNS/TNI/Polri/DPR,

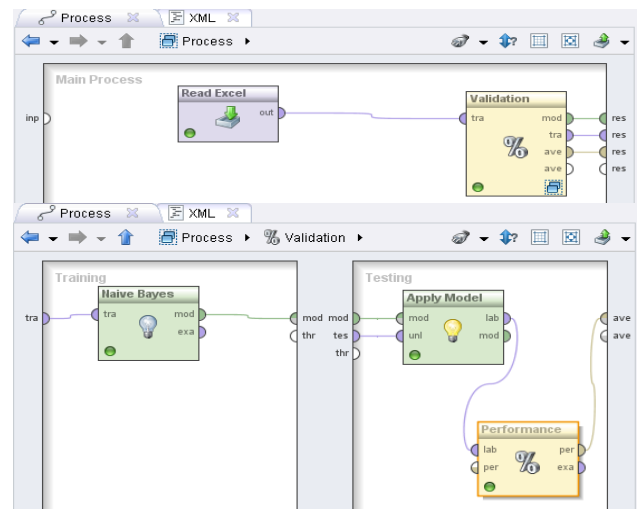
kepala rumah tangga tidak memiliki penyakit kronis, jumlah anggota keluarga 3, maka tidak mendapatkan KJP

17. Jika status kepemilikan rumah yaitu milik sendiri, anggota keluarga tidak ada yang usianya >65tahun, daya listrik sebesar 900, tidak bekerja sebagai PNS/TNI/Polri/DPR, kepala rumah tangga tidak memiliki penyakit kronis, jumlah anggota keluarga 4, dan pendidikan tertinggi anggota rumah tangga S1, maka tidak mendapatkan KJP
18. Jika status kepemilikan rumah yaitu milik sendiri, anggota keluarga tidak ada yang usianya >65tahun, daya listrik sebesar 900, tidak bekerja sebagai PNS/TNI/Polri/DPR, kepala rumah tangga tidak memiliki penyakit kronis, jumlah anggota keluarga 4, dan pendidikan tertinggi anggota rumah tangga SMA, maka mendapatkan KJP
19. Jika status kepemilikan rumah yaitu milik sendiri, anggota keluarga tidak ada yang usianya >65tahun, daya listrik sebesar 900, tidak bekerja sebagai PNS/TNI/Polri/DPR, kepala rumah tangga tidak memiliki penyakit kronis, jumlah anggota keluarga 5, maka mendapatkan KJP
20. Jika status kepemilikan rumah yaitu milik sendiri, anggota keluarga tidak ada yang usianya >65tahun, daya listrik sebesar 900, tidak bekerja sebagai PNS/TNI/Polri/DPR, kepala rumah tangga memiliki penyakit kronis, maka mendapatkan KJP
21. Jika status kepemilikan rumah yaitu milik sendiri, anggota keluarga tidak ada yang usianya >65tahun, daya listrik sebesar 900, bekerja sebagai PNS/TNI/Polri/DPR, maka tidak mendapatkan KJP
22. Jika status kepemilikan rumah yaitu milik sendiri, anggota keluarga tidak ada yang usianya >65tahun, maka mendapatkan KJP

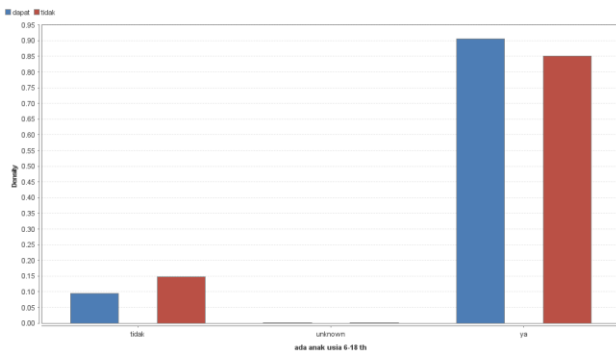
Dari hasil pohon keputusan dapat diketahui bahwa status kepemilikan rumah memiliki peran terbesar dalam keputusan penerimaan KJP karena berada di posisi yang teratas dalam diagram. Diikuti daya listrik tempat tinggal calon penerima, kondisi kepala keluarga dan pekerjaan orang tua.

Pengolahan Naive Bayes

Proses pengolahan *naive bayes* dengan membuat model desainnya terlebih dahulu, seperti dibawah ini:

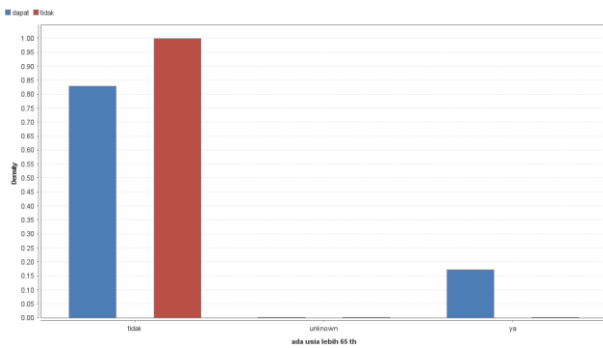


Gambar 5. Model Desain Menggunakan Naive Bayes



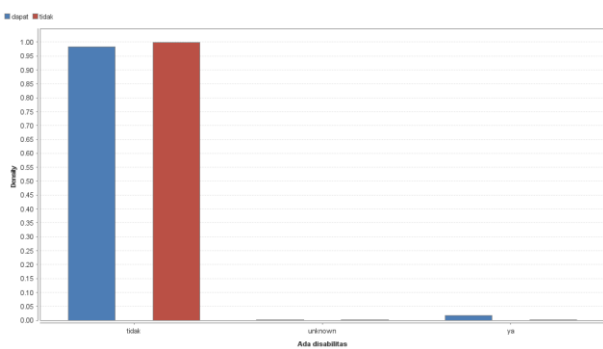
Gambar 12. Distribusi Data Pada Atribut Terdapat Anak Usia 6-18 Tahun

Dari gambar 12. grafik sebaran menurut terdapat anak usia 6-18 tahun atau tidak, secara garis besar untuk yang terdapat anak yang usia 6-18 tahun, yang tidak mendapatkan kartu KJP lebih sedikit dibandingkan yang mendapatkan kartu KJP dan untuk yang tidak terdapat anak usia 6-18 tahun yang tidak mendapatkan KJP lebih besar dibandingkan yang mendapatkan KJP.



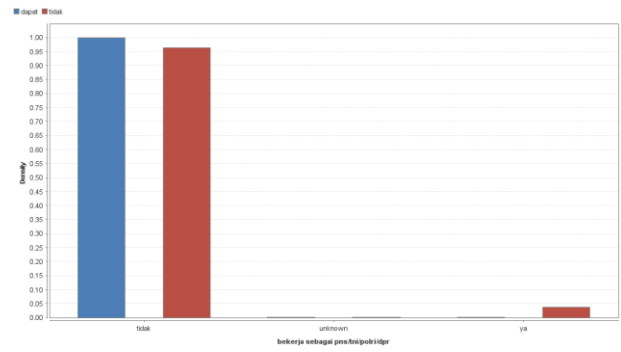
Gambar 13. Distribusi Data Pada Atribut Terdapat Anggota Keluarga Usia Lebih 65 Tahun

Dari gambar 13. grafik sebaran menurut terdapat anggota keluarga berusia lebih 65 tahun atau tidak, secara garis besar untuk yang terdapat anggota keluarga berusia 65 tahun, yang tidak mendapatkan kartu KJP lebih sedikit dibandingkan yang mendapatkan kartu KJP dan untuk yang tidak terdapat anggota keluarga usia lebih 65 tahun yang tidak mendapatkan KJP lebih besar dibandingkan yang mendapatkan KJP.



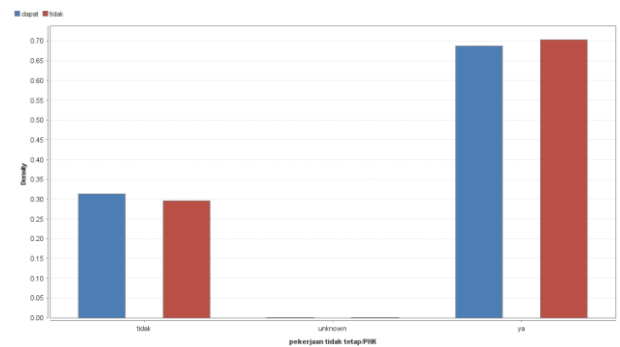
Gambar 14. Distribusi Data Pada Atribut Terdapat Disabilitas

Dari gambar 14. grafik sebaran menurut terdapat anggota keluarga disabilitas atau tidak, secara garis besar untuk yang terdapat anggota keluarga disabilitas, yang tidak mendapatkan kartu KJP lebih sedikit dibandingkan yang mendapatkan kartu KJP dan untuk yang tidak terdapat anggota keluarga disabilitas yang tidak mendapatkan KJP lebih besar dibandingkan yang mendapatkan KJP.



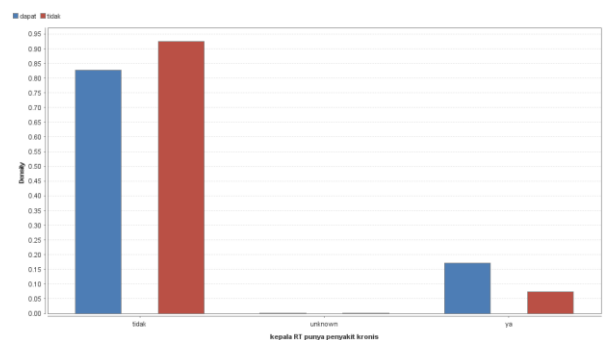
Gambar 15. Distribusi Data Pada Atribut Bekerja Sebagai PNS/TNI/Polri/DPR

Dari gambar 15. grafik sebaran menurut atribut bekerja sebagai PNS/TNI/Polri/DPR, secara garis besar untuk yang terdapat anggota keluarga bekerja sebagai PNS/TNI/Polri/DPR, yang tidak mendapatkan kartu KJP lebih banyak dibandingkan yang mendapatkan kartu KJP dan untuk yang tidak terdapat anggota keluarga bekerja sebagai PNS/TNI/Polri/DPR yang tidak mendapatkan KJP lebih sedikit dibandingkan yang mendapatkan KJP.



Gambar 16. Distribusi Data Pada Atribut Terdapat Anggota Keluarga Bekerja Tidak Tetap/PHK

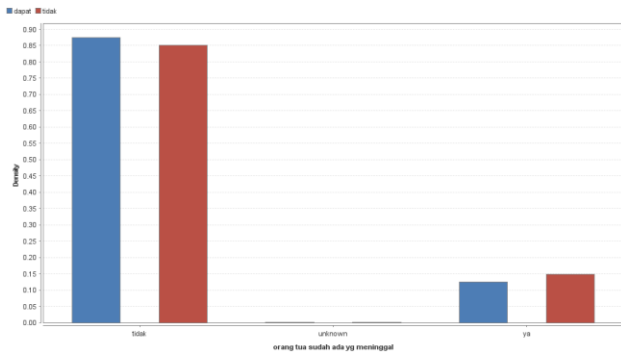
Dari gambar 16. grafik sebaran menurut atribut anggota keluarga bekerja tidak tetap/PHK, secara garis besar untuk yang terdapat anggota keluarga bekerja tidak tetap/PHK, yang tidak mendapatkan kartu KJP lebih besar dibandingkan yang mendapatkan kartu KJP dan untuk yang tidak terdapat anggota keluarga usia lebih 65 tahun yang tidak mendapatkan KJP lebih sedikit dibandingkan yang mendapatkan KJP.



Gambar 17. Distribusi Data Pada Atribut Kepala Rumah Tangga Mempunyai Penyakit Kronis

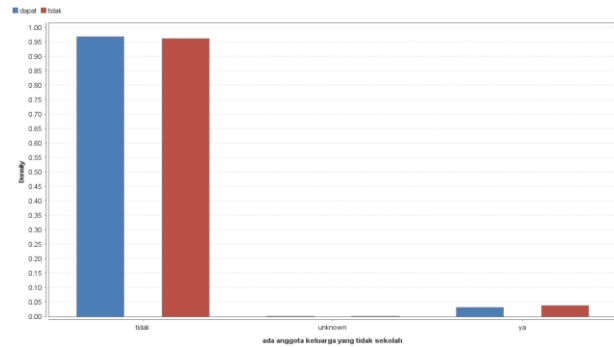
Dari gambar 17. grafik sebaran menurut atribut kepala rumah tangga punya penyakit kronis, secara garis besar untuk yang atribut kepala rumah tangga punya penyakit kronis, yang tidak mendapatkan kartu KJP lebih sedikit dibandingkan yang mendapatkan kartu KJP dan untuk yang tidak terdapat atribut kepala rumah tangga punya penyakit kronis yang tidak

mendapatkan KJP lebih besar dibandingkan yang mendapatkan KJP.



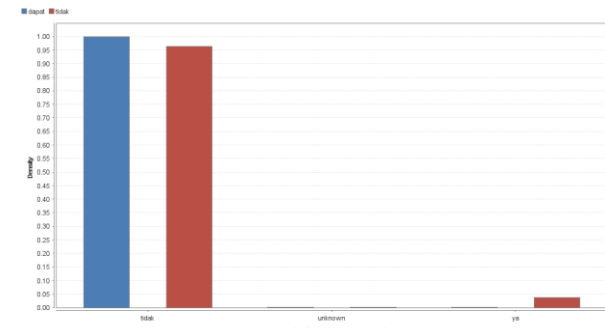
Gambar 18. Distribusi Data Pada Atribut Orang Tua Sudah Ada Yang Meninggal

Dari gambar 18. grafik sebaran menurut atribut orang tua sudah ada yang meninggal, secara garis besar untuk yang orang tua sudah ada yang meninggal, yang tidak mendapatkan kartu KJP lebih besar dibandingkan yang mendapatkan kartu KJP dan untuk yang tidak terdapat orang tua sudah ada yang meninggal yang tidak mendapatkan KJP lebih sedikit dibandingkan yang mendapatkan KJP.



Gambar 19. Distribusi Data Pada Atribut Terdapat Anggota Keluarga Yang Tidak Sekolah

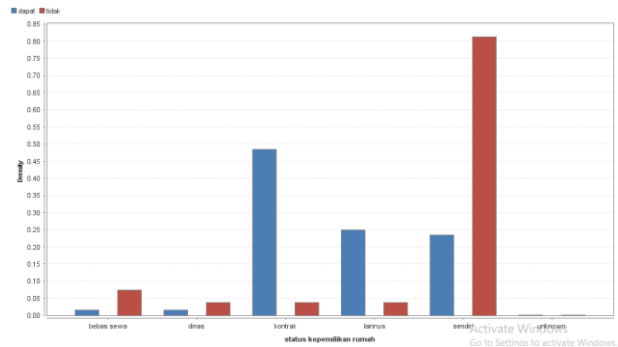
Dari gambar 19. grafik sebaran menurut atribut terdapat anggota keluarga yang tidak sekolah, secara garis besar untuk terdapat anggota keluarga yang tidak sekolah, yang tidak mendapatkan kartu KJP lebih besar dibandingkan yang mendapatkan kartu KJP dan untuk yang tidak terdapat anggota keluarga yang tidak sekolah yang tidak mendapatkan KJP lebih sedikit dibandingkan yang mendapatkan KJP.



Gambar 20. Distribusi Data Pada Atribut Penghasilan Orang Tua Mencukupi

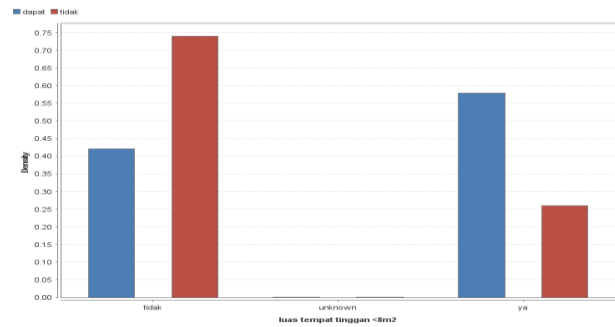
Dari gambar 20. grafik sebaran menurut atribut penghasilan orang tua mencukupi, secara garis besar untuk penghasilan orang tua mencukupi, yang tidak mendapatkan kartu KJP lebih besar dibandingkan yang mendapatkan kartu KJP dan untuk yang

penghasilan orang tua tidak mencukupi yang tidak mendapatkan KJP lebih sedikit dibandingkan yang mendapatkan KJP.



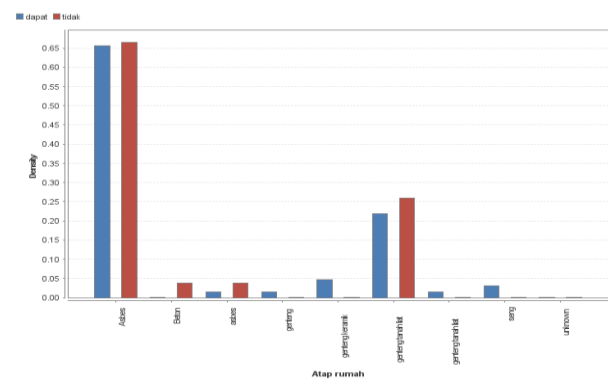
Gambar 21. Distribusi Data Pada Atribut Status Kepemilikan Rumah

Dari gambar 21. grafik sebaran menurut atribut status kepemilikan rumah, secara garis besar untuk status kepemilikan rumah dinas, bebas sewa dan milik sendiri, yang tidak mendapatkan kartu KJP lebih besar dibandingkan yang mendapatkan kartu KJP dan untuk yang status pemilikan kontrak atau lainnya yang tidak mendapatkan KJP lebih sedikit dibandingkan yang mendapatkan KJP.



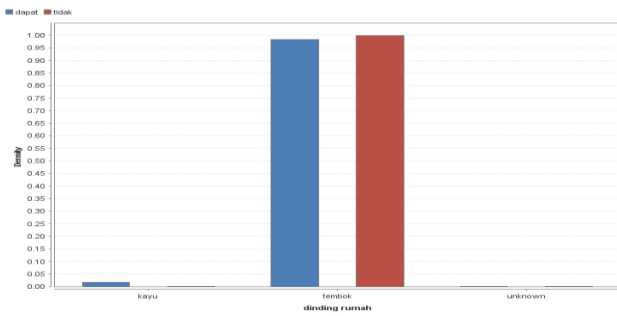
Gambar 22. Distribusi Data Pada Atribut Luas Tempat Tinggal <8m²

Dari gambar 22. grafik sebaran menurut atribut luas tempat tinggal <8m², secara garis besar untuk luas tempat tinggal <8m², yang tidak mendapatkan kartu KJP lebih sedikit dibandingkan yang mendapatkan kartu KJP dan untuk yang luas tempat tinggal tidak <8m² yang tidak mendapatkan KJP lebih besar dibandingkan yang mendapatkan KJP.



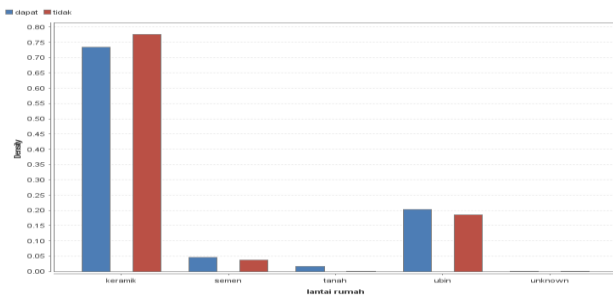
Gambar 23. Distribusi Data Pada Jenis Atap Rumah

Dari gambar 23. grafik sebaran menurut atribut jenis atap rumah, secara garis besar untuk asbes, genteng tanah liat dan beton yang tidak mendapatkan kartu KJP lebih besar dibandingkan yang mendapatkan kartu KJP dan genteng keramik, genteng tanah liat, dan seng yang tidak mendapatkan KJP lebih sedikit dibandingkan yang mendapatkan KJP.



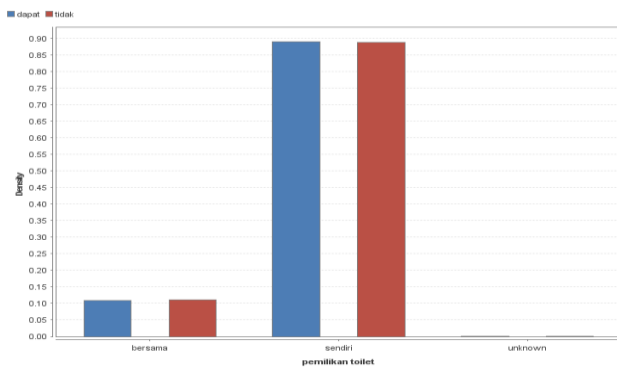
Gambar 24. Distribusi Data Pada Atribut Dinding Rumah

Dari gambar 24. grafik sebaran menurut atribut dinding rumah, secara garis besar untuk tembok yang tidak mendapatkan kartu KJP lebih besar dibandingkan yang mendapatkan kartu KJP dan dinding kayu yang tidak mendapatkan KJP lebih sedikit dibandingkan yang mendapatkan KJP.



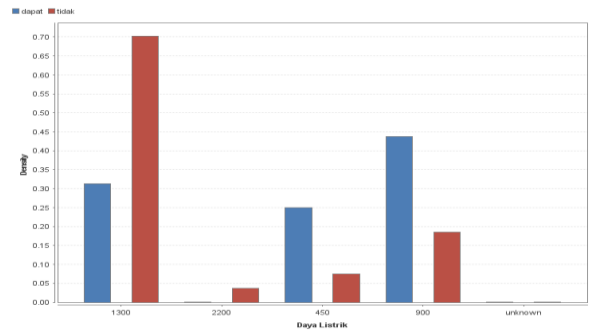
Gambar 25. Distribusi Data Pada Atribut Lantai Rumah

Dari gambar 25. grafik sebaran menurut atribut lantai rumah, secara garis besar untuk keramik yang tidak mendapatkan kartu KJP lebih besar dibandingkan yang mendapatkan kartu KJP sedangkan pada semen, tanah dan ubin yang tidak mendapatkan KJP lebih sedikit dibandingkan yang mendapatkan KJP.



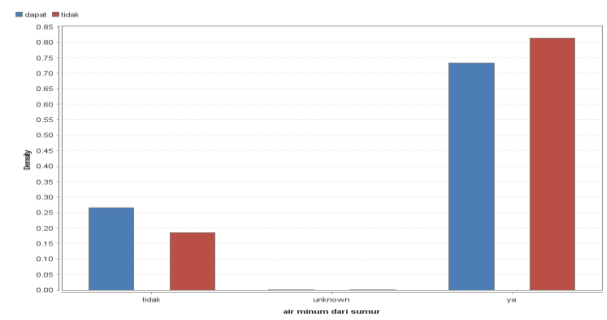
Gambar 26. Distribusi Data Pada Atribut Pemilikan Toilet

Dari gambar 26. grafik sebaran menurut atribut pemilikan toilet, secara garis besar untuk pemakaian bersama yang tidak mendapatkan kartu KJP sama dengan yang mendapatkan kartu KJP sedangkan pada pemakaian sendiri, yang tidak mendapatkan KJP lebih sedikit dibandingkan yang mendapatkan KJP.



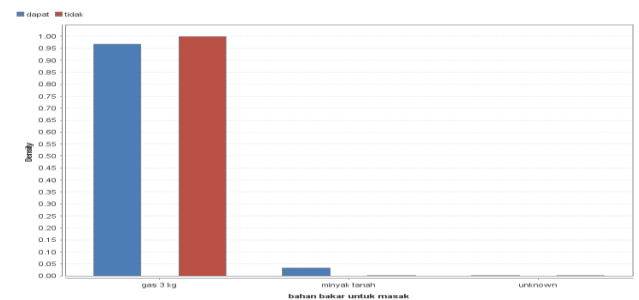
Gambar 27. Distribusi Data Pada Atribut Pemakaian Daya Listrik

Dari gambar 27. grafik sebaran menurut atribut pemakaian daya listrik, secara garis besar untuk pemakaian daya 1300 dan 2200 watt yang tidak mendapatkan kartu KJP lebih besar dengan yang mendapatkan kartu KJP sedangkan pada 450 dan 900 watt, yang tidak mendapatkan KJP lebih sedikit dibandingkan yang mendapatkan KJP.



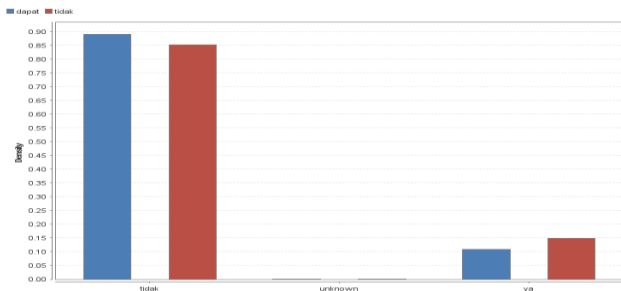
Gambar 28. Distribusi Data Pada Atribut Air Minum Dari Sumur

Dari gambar 28. grafik sebaran menurut atribut air minum dari sumur, secara garis besar untuk yang memakai air minum dari sumur yang tidak mendapatkan kartu KJP lebih besar dengan yang mendapatkan kartu KJP sedangkan yang tidak minum dari air sumur, yang tidak mendapatkan KJP lebih sedikit dibandingkan yang mendapatkan KJP.



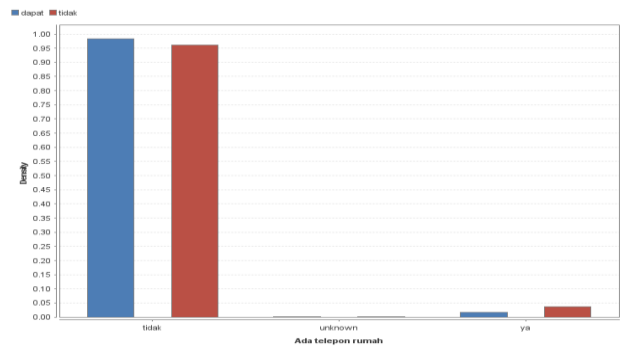
Gambar 29. Distribusi Data Pada Atribut Bahan Bakar Untuk Memasak

Dari gambar 29. grafik sebaran menurut atribut bahan bakar untuk memasak, secara garis besar untuk yang memakai gas 3 kg yang tidak mendapatkan kartu KJP lebih besar dengan yang mendapatkan kartu KJP sedangkan yang memakai minyak tanah, yang tidak mendapatkan KJP lebih sedikit dibandingkan yang mendapatkan KJP.



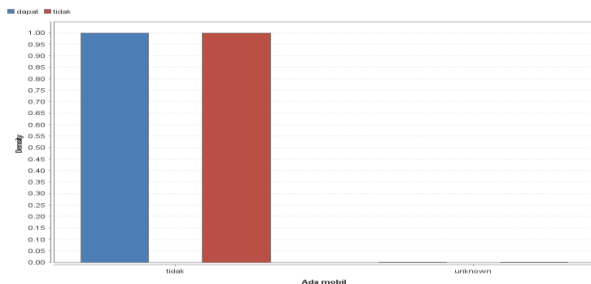
Gambar 30. Distribusi Data Pada Atribut Perubahan Kondisi Sosial Ekonomi Dibanding Tahun Sebelumnya

Dari gambar 30. grafik sebaran menurut atribut perubahan kondisi sosial ekonomi dibanding tahun sebelumnya, secara garis besar untuk yang terdapat perubahan kondisi sosial ekonomi dibanding tahun sebelumnya yang tidak mendapatkan kartu KJP lebih besar dengan yang mendapatkan kartu KJP sedangkan yang tidak terdapat perubahan kondisi sosial ekonomi dibanding tahun sebelumnya, yang tidak mendapatkan KJP lebih sedikit dibandingkan yang mendapatkan KJP.



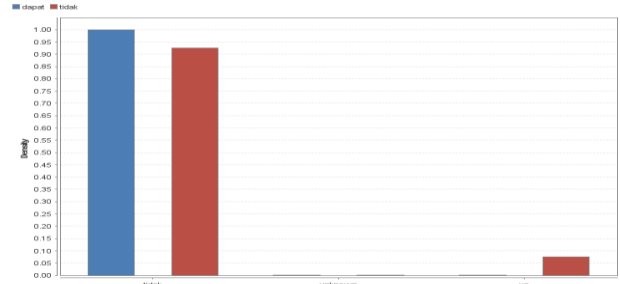
Gambar 33. Distribusi Data Pada Atribut Mempunyai Telepon Rumah

Dari gambar 33. grafik sebaran menurut atribut mempunyai telepon rumah, secara garis besar untuk yang mempunyai telepon rumah yang tidak mendapatkan kartu KJP lebih besar dibanding yang mendapatkan kartu KJP sedangkan yang tidak mempunyai telepon rumah, yang tidak mendapatkan KJP lebih sedikit dibandingkan dengan yang mendapatkan KJP.



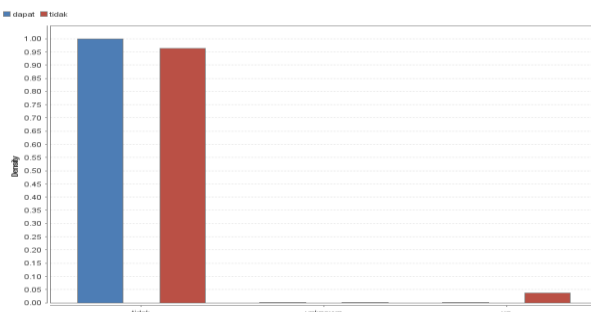
Gambar 31. Distribusi Data Pada Atribut Mempunyai Mobil

Dari gambar 31. grafik sebaran menurut atribut mempunyai mobil, secara garis besar untuk yang mempunyai mobil tidak mendapatkan kartu KJP sedangkan yang tidak mempunyai mobil, yang tidak mendapatkan KJP sebanding dengan yang mendapatkan KJP.



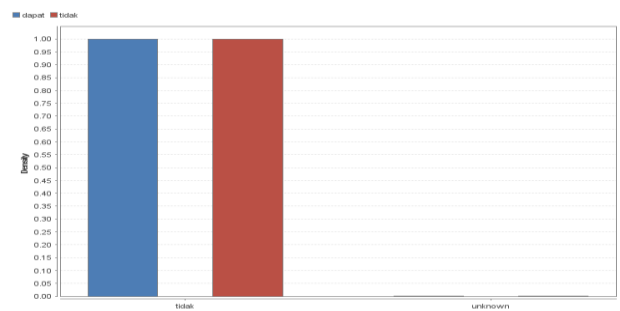
Gambar 34. Distribusi Data Pada Atribut Mempunyai AC

Dari gambar 34. grafik sebaran menurut atribut mempunyai AC, secara garis besar untuk yang mempunyai AC yang tidak mendapatkan kartu KJP lebih besar dibanding yang mendapatkan kartu KJP sedangkan yang tidak mempunyai AC, yang tidak mendapatkan KJP lebih sedikit dibandingkan dengan yang mendapatkan KJP.



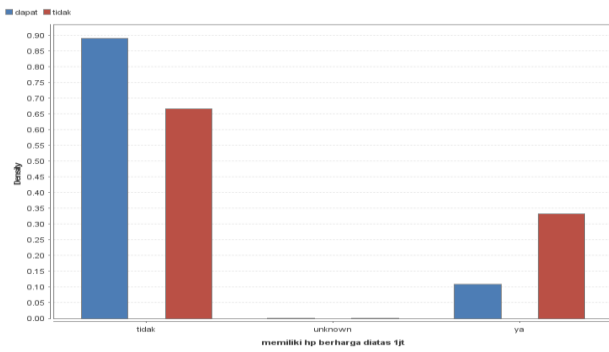
Gambar 32. Distribusi Data Pada Atribut Mempunyai Motor >2

Dari gambar 32. grafik sebaran menurut atribut mempunyai motor >2, secara garis besar untuk yang mempunyai motor >2 yang tidak mendapatkan kartu KJP lebih besar dibanding yang mendapatkan kartu KJP sedangkan yang tidak mempunyai motor >2, yang tidak mendapatkan KJP lebih sedikit dibandingkan dengan yang mendapatkan KJP.



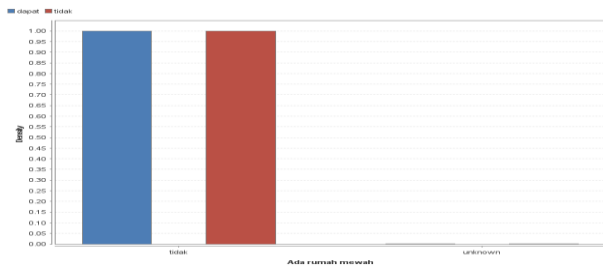
Gambar 35. Distribusi Data Pada Atribut Memakai Tabung Gas > 5,5kg

Dari gambar 35. grafik sebaran menurut atribut memakai tabung gas >5,5kg, secara garis besar untuk memakai tabung gas >5,5kg tidak mendapatkan kartu KJP sedangkan yang tidak memakai tabung gas >5,5kg, yang tidak mendapatkan KJP sama besarnya dengan yang mendapatkan KJP.



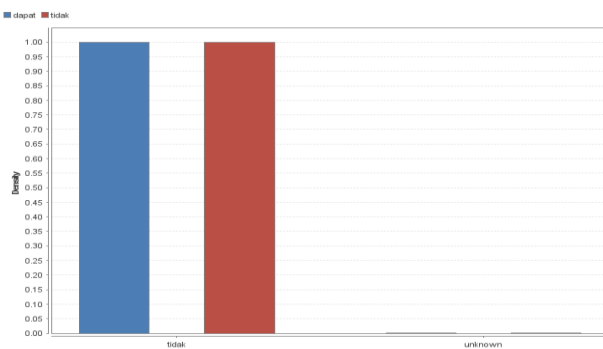
Gambar 36. Distribusi Data Pada Atribut Memiliki HP Berharga Diatas 1Juta

Dari gambar 36. grafik sebaran menurut atribut memiliki HP berharga diatas 1 juta, secara garis besar yang memiliki HP berharga diatas 1 juta yang tidak mendapatkan kartu KJP lebih besar dibandingkan yang mendapatkan kartu KJP sedangkan yang tidak memiliki HP berharga diatas 1 juta, yang tidak mendapatkan KJP lebih sedikit dibandingkan dengan yang mendapatkan KJP.



Gambar 37. Distribusi Data Pada Atribut Memiliki Rumah Mewah

Dari gambar 37. grafik sebaran menurut atribut memiliki rumah mewah, secara garis besar untuk yang memiliki rumah mewah tidak mendapatkan kartu KJP sedangkan yang tidak memiliki rumah mewah, yang tidak mendapatkan KJP sama besarnya dengan yang mendapatkan KJP.



Gambar 38. Distribusi Data Pada Atribut Memiliki Usaha Rumah Kontrakan

Dari gambar 38. grafik sebaran menurut atribut memiliki usaha rumah kontrakan, secara garis besar untuk yang memiliki usaha rumah kontrakan tidak mendapatkan kartu KJP sedangkan yang tidak memiliki usaha rumah kontrakan, yang tidak mendapatkan KJP sama besarnya dengan yang mendapatkan KJP.

Performance ID3

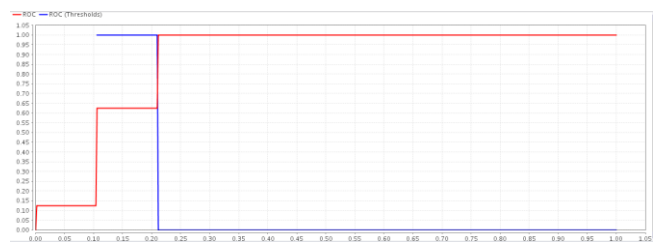
Berikut tingkat akursi menggunakan *confussion matrix* pada ID3:

accuracy: 77.78%			
	true dapat	true tidak	class precision
pred. dapat	17	4	80.95%
pred. tidak	2	4	66.67%
class recall	89.47%	50.00%	

Gambar 39. tingkat akursi menggunakan *confussion matrix* pada ID3

Gambar 39 adalah hasil akurasi data menggunakan ID3 sebesar 77,78%. Diketahui dari 91 data didapat 17 data prediksi *yes* sesuai real dengan *yes*, 4 prediksi *yes* ternyata *no*, 2 data prediksi *no* ternyata *yes*, dan 4 data predisi *no* sesuai dengan *no*. Dengan persentase prediksi yang mendapatkan KJP sebesar 80,95% dan persentase prediksi yang tidak mendapatkan KJP sebesar 66,67%. Persentase *class recall* pada yang mendapatkan KJP sebesar 89,47% dan persentase *class recall* yang benar tidak mendapatkan KJP sebesar 50,00%.

Hasil visualisasi kurva ROC pada ID3 dapat dilihat pada gambar 40. Garis horizontal merupakan *false positive* dan garis vertical merupakan *true positive*.



Gambar 40. Nilai Grafik ROC Pada ID3

Pada gambar 40. Nilai grafik ROC sebesar 0,868 sehingga dapat dinyatakan *Good Classification*. Karena nilai akurasi berada pada rentang 0.80-0.90.

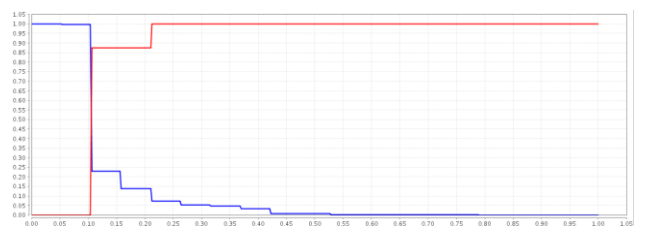
Performance Naive Bayes

Tingkat akursi menggunakan *confussion matrix* pada *naive bayes* terlihat pada gambar 41:

accuracy: 66.67%			
	true dapat	true tidak	class precision
pred. dapat	17	7	70.83%
pred. tidak	2	1	33.33%
class recall	89.47%	12.50%	

Gambar 41. tingkat akursi menggunakan *confussion matrix* pada *naive bayes*

Gambar 41 adalah hasil akurasi data menggunakan *naive bayes* sebesar 66,67%. Diketahui dari 91 data didapat 17 data prediksi *yes* sesuai real dengan *yes*, 7 prediksi *yes* ternyata *no*, 2 data prediksi *no* ternyata *yes*, dan 1 data predisi *no* sesuai dengan *no*. Dengan persentase prediksi yang mendapatkan KJP sebesar 70,83% dan persentase prediksi yang tidak mendapatkan KJP sebesar 33,33%. Persentase *class recall* pada yang mendapatkan KJP sebesar 89,47% dan persentase *class recall* yang benar tidak mendapatkan KJP sebesar 12,50%.



Gambar 42. Nilai Grafik ROC Pada ID3

Pada gambar 42. Nilai grafik ROC sebesar 0,882 sehingga dapat dinyatakan *Good Classification*. Karena nilai akurasi berada pada rentang 0.80-0.90

KESIMPULAN

Berdasarkan pembahasan yang telah diuraikan didapatkan penggunaan klasifikasi Analisa Kartu Jakarta Pintar dapat dilakukan dengan menggunakan algoritma ID3 dan Naïve Bayes. Algoritma ini dapat disimpulkan bahwa kriteria kepemilikan rumah memiliki peran terbesar dalam pengambilan keputusan diikuti jenis pekerjaan orang tua. Dimana pekerjaan orang tua buruh dan mempunyai tingkat pendidikan SMA kebawah memiliki peluang lebih mendapatkan KJP. Berdasarkan keakuratan algoritma ID3 memiliki 77,78% sedangkan klasifikasi Naïve Bayes memiliki tingkat keakuratan dalam prediksi 66,67%. Sehingga dapat disimpulkan berdasarkan tingkat akurasi bahwa permodelan ID3 memiliki keakuratan yang lebih tinggi. Sedangkan berdasarkan perhitungan ROC dalam mengukur akurasi dan kehandalan kedua algoritma berada di rentang *Good Classification* dimana grafik ROC algoritma ID3 yaitu 0.868 sedangkan Naïve Bayes 0.882.

REFERENSI

- [1] D. P. D. J. Pusat Data dan Teknologi Informasi Komunikasi Pendidikan, "Indikator Pendidikan 2020-2021."
- [2] M. R. Fahdia, N. Asnonik, and S. A. Putri, "IMPLEMENTASI METODE NAIVE BAYES PADA ANALISA PENERIMAAN KARTU JAKARTA PINTAR DI SMK AL KAUTSAR JAKARTA," *Inti Nusa Mandiri*, vol. 14, no. 2, pp. 231–236, 2020.
- [3] E. W. Ningsih and Hardiyanto, "Penerapan Algoritma Naïve Bayes Dalam Penentuan Kelayakan Penerima Kartu Jakarta Pintar Plus," vol. VI, no. 1, pp. 15–20, 2020.
- [4] M. Yunus, H. Ramadhan, D. R. Aji, and A. Yulianto, "Penerapan Metode Data Mining C4.5 Untuk Pemilihan Penerima Kartu Indonesia Pintar (KIP)," *Paradig. - J. Komput. dan Inform.*, vol. 23, no. 2, 2021.
- [5] N. Alfiah, "Klasifikasi Penerima Bantuan Sosial Program Keluarga Harapan Menggunakan Metode Naive Bayes," *J. Teknol. Inf.*, vol. XVI, pp. 32–40, 2021.
- [6] A. Saifudin, "Metode Data Mining Untuk Seleksi Calon Mahasiswa," *J. Teknol. Univ. Muhammadiyah Jakarta*, vol. 10, no. 1, pp. 25–36, 2018.
- [7] Suyanto, *Data Mining Untuk Klasifikasi Dan Klasterisasi Data*. Bandung: Informatika, 2019.
- [8] M. Arhami and M. Nasir, *DATA MINING-Algoritma dan Implementasi*, I. Yogyakarta: Andi Publishing, 2020.
- [9] Y. Mardi, "Data Mining : Klasifikasi Menggunakan Algoritma C4.5," *J. Edik Inform.*, vol. V2.i2 ISSN, pp. 213–219, 2017.
- [10] E. Fitriani, "Perbandingan Algoritma C4.5 Dan Naïve Bayes Untuk Menentukan Kelayakan Penerima Bantuan Program Keluarga Harapan," *Sistemasi*, vol. 9, no. 1, p. 103, 2020.