



**KLASIFIKASI KELAYAKAN PENERIMA KARTU JAKARTA
PINTAR (KJP) DENGAN SELEKSI FITUR BACKWARD
ELIMINATION MENGGUNAKAN ALGORITMA
KLASIFIKASI NAÏVE BAYES**

SKRIPSI

JAMALUL IKHSAN

1810511099

**UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAKARTA
PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
2022**



**KLASIFIKASI KELAYAKAN PENERIMA KARTU JAKARTA
PINTAR (KJP) DENGAN SELEKSI FITUR BACKWARD
ELIMINATION MENGGUNAKAN ALGORITMA
KLASIFIKASI NAÏVE BAYES**

SKRIPSI

**Diajukan Sebagai Salah Satu Syarat Untuk Memperoleh Gelar Sarjana
Komputer**

JAMALUL IKHSAN

1810511099

**UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAKARTA
PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER**

2022

PERNYATAAN ORISINALITAS

Dengan ini dinyatakan bahwa Skripsi Tugas Akhir ini adalah hasil karya saya sendiri, dan semua sumber yang dikutip maupun yang dirujuk dalam penelitian skripsi ini telah saya nyatakan dengan benar.

Nama : Jamalul Ikhsan

NIM : 1810511099

Tanggal : 23 Juni 2022

Bilamana dikemudian hari ditemukan ketidaksesuaian dengan pernyataan saya ini, maka saya bersedia dituntut dan diproses sesuai dengan ketentuan yang berlaku.

Jakarta, 23 Juni 2022

Yang Menyatakan,



(Jamalul Ikhsan)

PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI UNTUK KEPENTINGAN AKADEMIS

Sebagai civitas akademis Universitas Pembangunan Nasional Veteran Jakarta saya yang bertanda tangan dibawah ini :

Nama : Jamalul Ikhsan

NIM : 1810511099

Fakultas : Ilmu Komputer

Program Studi : Informatika

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Pembangunan Nasional Veteran Jakarta Hak Bebas Royalti Non eksklusif (Non-exclusive Royalti Free Right) atas karya ilmiah saya yang berjudul :

“Klasifikasi Kelayakan Pengguna Kartu Jakarta Pintar Dengan Seleksi Fitur *Backward Elimination* Menggunakan Algoritma Klasifikasi *Naïve Bayes*”

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti ini Universitas Pembangunan Nasional Veteran Jakarta berhak menyimpan, mengalih media/formatkan, mengelola dalam bentuk pangkalan data (databases), merawat, dan mempublikasikan Skripsi saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat oleh : Jamalul Ikhsan

Waktu dibuat : Jakarta, 23 Juni 2022

Yang Menyatakan,



(Jamalul Ikhsan)

LEMBAR PERSETUJUAN

Dengan ini dinyatakan bahwa Skripsi Tugas Akhir berikut :

Nama : Jamalul Ikhsan

NIM : 1810511099

Program Studi : Informatika

Judul : Klasifikasi Kelayakan Penerima Kartu Jakarta Pintar Dengan Seleksi Fitur *Backward Elimination* Menggunakan Algoritma Klasifikasi *Naïve Bayes*.

Sebagai bagian dari persyaratan yang diperlukan untuk mengikuti sidang Skripsi pada Program Studi Informatika Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Veteran Jakarta.

Menyetujui,

Dosen Pembimbing I



(Catur Nugrahaeni., M.Kom.)

Menyetujui,

Dosen Pembimbing II



(Desta Sandya Prasvita, S.Komp., M.Kom)

Menyetujui,

Ketua Program Studi



(Desta Sandya Prasvita, S.Komp., M.Kom)

Ditetapkan di : Jakarta

Tanggal ditetapkan : 23 Juni 2022

PENGESAHAN

Dengan ini menyatakan bahwa Skripsi berikut ini;

Nama : Jamalul Ikhsan

NIM : 1810511099

Program Studi : SI Informatika

Judul :Klasifikasi Kelayakan Penerima Kartu Jakarta Pintar
Dengan Seleksi Fitur *Backward Elimination* Menggunakan
Algoritma Klasifikasi *Naïve Bayes*.

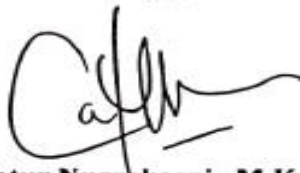
Telah berhasil dipertahankan di hadapan Tim penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Komputer pada Program Studi Informatika, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Veteran Jakarta.



(Nur Hafifah Matondang, S.kom,
MM., M.T.I.)
Penguji I



(Nurul Chaudidah S.Kom., M.kom.)
Penguji II



(Catur Nugrahaeni., M.Kom.)
Pembimbing I



(Desta Sandya Prasvita S.Komp., M.Kom)
Pembimbing II



(Dr. Ermatita., M.Kom.)
Dekan



(Desta Sandya Prasvita S.Komp., M.Kom)
Ketua Program Studi

Ditetapkan di : Jakarta
Tanggal Ujian : 13 Juli 2022



Klasifikasi Kelayakan Penerima Kartu Jakarta Pintar Dengan Seleksi Fitur *Backward Elimination* Menggunakan Algoritma Klasifikasi *Naïve Bayes*.

Jamalul Ikhsan

ABSTRAK

Kartu Jakarta Pintar atau biasa disebut sebagai KJP merupakan program yang diciptakan oleh pemerintah provinsi DKI Jakarta untuk memberikan bantuan pendidikan di wilayah DKI Jakarta, bantuan program pendidikan ini diharapkan dapat mengurangi tingkat putus sekolah yang terjadi dikalangan masyarakat kelas menengah kebawah. Tentunya setiap program memiliki sistem seleksi seperti halnya pada program Kartu Jakarta Pintar ini, pemerintah melakukan proses seleksi untuk memastikan masyarakat yang pantas untuk mendapatkan bantuan dan yang tidak. Dalam proses seleksi penerimaan terkadang terdapat suatu kelalaian yang dilakukan oleh pihak operator dalam melakukan input data sehingga yang seharusnya masyarakat bisa mendapatkan bantuan tetapi karena adanya kelalaian tersebut menjadi tidak dapat bantuan, selain itu adanya proses seleksi yang bersifat subyektif sehingga bantuan yang diberikan oleh pemerintah ini sering tidak tepat sasaran. Maka dari itu, penelitian ini bertujuan untuk membuat suatu sistem yang dapat melakukan langsung menentukan kelayakan penerima kartu jakarta pintar dengan tepat dan akurat menggunakan metode algoritma klasifikasi naïve bayes dan seleksi fitur *backward elimination* untuk menentukan fitur yang berperan penting dalam penentuan kelayakan penerima kartu jakarta pintar. Dataset didapatkan dari pihak sekolah dengan jumlah sebesar 158 data calon peserta penerima. Berdasarkan hasil penelitian ini, bahwa akurasi yang didapatkan oleh model dan setelah dilakukan perhitungan akurasi menggunakan confusion matrix terhadap model diperoleh dengan akurasi sebesar 90.625%, namun dengan bantuan seleksi fitur *backward elimination*, akurasi dapat meningkat cukup signifikan menjadi 96.875%.

Kata kunci : Klasifikasi, Kartu Jakarta Pintar, Naïve Bayes, Machine Learning, Data Mining.

Eligibility Classification of Jakarta Smart Card Recipients With Backward Elimination Feature Selection Using Naïve Bayes Classification Algorithm.

Jamalul Ikhsan

ABSTRACT

The Jakarta Smart Card or commonly referred to as KJP is a program created by the provincial government of DKI Jakarta to provide educational assistance in the DKI Jakarta area. Of course, every program has a selection system, as in the Jakarta Smart Card program, the government conducts a selection process to ensure that people who deserve assistance and who don't. In the acceptance selection process sometimes there is an omission by the operator in inputting data so that the community should be able to get help but because of this negligence they cannot get help, besides that there is a subjective selection process so that the assistance provided by the government is often not on target. Therefore, this study aims to create a system that can directly and accurately determine the eligibility of Jakarta smart card recipients using the naive Bayes classification algorithm and backward elimination feature selection to determine the features that play an important role in determining the eligibility of Jakarta smart card recipients. The dataset was obtained from the school with a total of 158 data on prospective recipients. Based on the results of this study, the accuracy obtained by the model and after calculating the accuracy using the confusion matrix on the model is obtained with an accuracy of 90.625%, but with the help of backward elimination feature selection, the accuracy can increase significantly to 96.875%.

Keyword : *Classification, Jakarta Smart Card, Naïve Bayes, Machine Learning, Data Mining.*

KATA PENGANTAR

Puji Syukur penulis panjatkan atas kehadiran dan keberkahan Allah SWT dan segala karunia-Nya sehingga Skripsi dengan judul “Klasifikasi Kelayakan Penerima Kartu Jakarta Pintar Dengan Seleksi Fitur *Backward Elimination* Menggunakan Algoritma Klasifikasi *Naïve Bayes*.” telah berhasil untuk diselesaikan. Skripsi ini disusun untuk mendapatkan gelar sarjana computer pada Program Studi Informatika Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Veteran Jakarta.

Disamping itu, dalam proses penulisan dan penyusunan skripsi ini penulis mendapatkan banyak bantuan, baik itu dukungan secara moral maupun materil dari berbagai pihak. Oleh karena itu, penulis ingin mengucapkan terima kasih kepada :

1. Ibu Catur Nugrahaeni., M.Kom. selaku dosen pembimbing pertama yang telah banyak memberikan saran dalam proses penyusunan skripsi ini.
2. Bapak Desta Sandya Prasvita, S.Komp., M.Kom selaku dosen pembimbing kedua yang juga sangat membantu dalam memberikan arahan, dan saran dalam penyelesaian penyusunan skripsi ini dari awal hingga akhir.
3. Teman-teman program studi informatika dan salah satunya yang sangat berjasa yaitu Kak Raihan Alif yang memberikan arahan dan bimbingan nya mulai dari awal hingga akhir penyusunan skripsi ini.
4. Kedua orang tua saya yang selalu memberikan dukungan dan doa atas keberhasilan dalam penyusunan skripsi ini.
5. Semua pihak yang tidak dapat disebutkan satu-persatu nama nya, yang telah juga memberikan kontribusi terhadap penyusunan skripsi ini baik secara langsung atau tidak langsung.

Penulis menyadari Skripsi ini tidak lepas dari kesalahan dan kekurangan oleh karena itu, penulis bersedia menerima kritik dan saran yang dapat membangun untuk melakukan perbaikan diri. Penulis berharap semoga Skripsi ini dapat memberikan manfaat.

Jakarta, 23 Juni 2022

Penulis.

DAFTAR ISI

HALAMAN JUDUL	i
PERNYATAAN ORISINALITAS	iii
PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI UNTUK KEPENTINGAN AKADEMIS	iv
LEMBAR PERSETUJUAN	v
ABSTRAK	vii
<i>ABSTRACT</i>	viii
KATA PENGANTAR.....	ix
DAFTAR ISI.....	x
DAFTAR GAMBAR.....	xii
DAFTAR TABEL	xiii
BAB 1	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	4
1.3 Tujuan Penelitian	4
1.4 Manfaat Penelitian.....	5
1.5 Batasan Masalah.....	5
1.6 Luaran Yang Diharapkan	5
1.7 Sistematika Penulisan	6
BAB 2	7
2.1 Kartu Jakarta Pintar	7
2.2 Data Mining	7
2.2.1 Pengertian Data Mining.....	7
2.2.2 Tahapan Dalam Data Mining.....	8
2.1 Backward Elimination.....	10
2.2 Machine Learning.....	11
2.3 Klasifikasi	12
2.4 Algoritma Naïve Baiyes	12
2.5 Confusion Matrix.....	14
2.6 Studi Literatur	16
BAB 3	20
3.1 Kerangka Pikir	20

3.1.1	Identifikasi Masalah	21
3.1.2	Studi Literatur	21
3.1.3	Pengumpulan Data.....	21
3.1.4	Praproses Data	23
3.1.5	Pembagian Data	24
3.1.6	Data latih	25
3.1.7	Training.....	25
3.1.8	Model naïve bayes	25
3.1.9	Data uji.....	26
3.1.10	Testing	26
3.1.11	Seleksi Fitur Backward Elimination	26
3.1.12	Hasil Akurasi Pengujian	27
3.2	Perangkat Penelitian.....	28
3.3	Jadwal Penelitian	29
BAB 4	30
4.1	Data	30
4.2	Pra-Proses Data	33
4.3	Pembagian Data.....	37
4.4	Perancangan Model	38
4.5	Confusion Matrix.....	66
4.6	Backward Elimination.....	66
BAB 5	69
5.1	Kesimpulan	69
5.2	Saran	70
DAFTAR PUSTAKA	71
RIWAYAT HIDUP	74
LAMPIRAN	75

DAFTAR GAMBAR

Gambar 2.1 Tahapan Data Mining.....	8
Gambar 2.2 Flowchart Seleksi Fitur Backward Elimination.....	10
Gambar 3.1 Kerangka Pikir.....	36

DAFTAR TABEL

Tabel 2.1 Tabel Confusion Matrix.....	14
Tabel 3.1 Fitur Dataset.....	22
Tabel 3.2 Tabel Confusion Matrix.....	27
Tabel 3.3 Perangkat Keras Penelitian	28
Tabel 3.4 Perangkat Lunak Penelitian	29
Tabel 3.5 Jadwal Penelitian.....	29
Tabel 4.1 Variabel, Tipe Data, dan Atribut.	30
Tabel 4.2 Informasi Data.....	33
Tabel 4.3 Tabel Data Sebelum Proses Transformasi Data.	37
Tabel 4.4 Tabel Data Setelah Proses Transformasi Data.	37
Tabel 4.5 Tabel Data Sample untuk Training Data.....	39
Tabel 4.6 Tabel Probabilitas atribut kepala rumah tangga memiliki pendidikan SMA kebawah.....	40
Tabel 4.7 Tabel Probabilitas atribut pendidikan anggota rumah tangga yang sudah tidak sekolah.	40
Tabel 4.8 Tabel Probabilitas atribut apakah ada ibu hamil dalam anggota rumah tangga.	41
Tabel 4.9 Tabel Probabilitas atribut Apakah ada anak usia sekolah (6-18 tahun) dalam anggota rumah tangga.	41
Tabel 4.10 Tabel Probabilitas atribut apakah ada anggota lanjut usia (usia lebih dari 65 tahun) dalam rumah tangga.	41
Tabel 4.11 Tabel Probabilitas atribut apakah ada anggota disabilitas/penyandang cacat dalam rumah tangga.	41
Tabel 4.12 Tabel Probabilitas apakah ada anggota yang bekerja sebagai PNS/TNI/POLRI/DPR dalam rumah tangga.	41
Tabel 4.13 Tabel Probabilitas atribut Kepala Rumah Tangga mempunyai pekerjaan tidak tetap/PHK.	41
Tabel 4.14 Tabel Probabilitas atribut Kepala rumah tangga memiliki penyakit kronis yang menyulitkan untuk bekerja.	42
Tabel 4.15 Tabel Probabilitas atribut salah Satu/Kedua Orang Tua Kandung Meninggal Dunia.....	42
Tabel 4.16 Tabel Probabilitas atribut Ada anak umur 7-18 tahun yang tidak sekolah karena alasan ekonomi.	42
Tabel 4.17 Tabel Probabilitas atribut Ayah/ibu pegawai/karyawan swasta/wiraswatawan berpenghasilan cukup.....	42
Tabel 4.18 Tabel Probabilitas atribut Ayah/ibu pegawai/karyawan swasta/wiraswatawan berpenghasilan cukup (kategori penghasilan).	42
Tabel 4.19 Tabel Probabilitas atribut jumlah kepala keluarga dalam rumah tangga.	43
Tabel 4.20 Tabel Probabilitas atribut Status kepemilikan Tempat Tinggal.	43
Tabel 4.21 Tabel Probabilitas atribut luas lantai bangunan tempat tinggal kurang dari 8 m ²	43
Tabel 4.22 Tabel Probabilitas atribut bahan bangunan utama atap rumah terluas.....	43
Tabel 4.23 Tabel Probabilitas atribut bahan bangunan utama dinding rumah terluas.	44
Tabel 4.24 Tabel Probabilitas atribut bahan bangunan utama lantai rumah terluas.....	44
Tabel 4.25 Tabel Probabilitas atribut penggunaan fasilitas tempat buang air besar.	44
Tabel 4.26 Tabel Probabilitas atribut sumber penerangan utama rumah tangga.	45

Tabel 4.27 Tabel Probabilitas atribut sumber penerangan utama rumah tangga (kategori daya listrik rumah).	45
Tabel 4.28 Tabel Probabilitas atribut Sumber air minum berasal dari sumur.	45
Tabel 4.29 Tabel Probabilitas atribut bahan bakar untuk memasak.	45
Tabel 4.30 Tabel Probabilitas atribut apakah terdapat perubahan kondisi sosial ekonomi rumah tangga dari tahun sebelumnya.....	46
Tabel 4.31 Tabel Probabilitas atribut apakah rumah tangga memiliki mobil.	46
Tabel 4.32 Tabel Probabilitas atribut memiliki kapal motor.	46
Tabel 4.33 Tabel Probabilitas atribut memiliki perahu motor lebih dari atau sama dengan dua.....	46
Tabel 4.34 Tabel Probabilitas atribut memiliki sepeda motor lebih dari dua.	46
Tabel 4.35 Tabel Probabilitas atribut memiliki telepon rumah (bukan hp).	47
Tabel 4.36 Tabel Probabilitas atribut apakah rumah tangga memiliki AC.	47
Tabel 4.37 Tabel Probabilitas atribut apakah rumah tangga memiliki tabung gas lebih dari 5,5 kg.....	47
Tabel 4.38 Tabel Probabilitas atribut Anak memiliki HP berharga lebih Rp1.000.000....	47
Tabel 4.39 Tabel Probabilitas atribut rumah mewah.	47
Tabel 4.40 Tabel Probabilitas atribut usaha rumah kontrakan atau kos.	47
Tabel 4.41 Tabel Probabilitas atribut status.	48
Tabel 4.42 Tabel Probabilitas atribut kepala rumah tangga memiliki pendidikan SMA kebawah.....	48
Tabel 4.43 Tabel Probabilitas atribut pendidikan anggota rumah tangga yang sudah tidak sekolah.	48
Tabel 4.44 Tabel Probabilitas atribut apakah ada ibu hamil dalam anggota rumah tangga.	49
Tabel 4.45 Tabel Probabilitas atribut Apakah ada anak usia sekolah (6-18 tahun) dalam anggota rumah tangga.	49
Tabel 4.46 Tabel Probabilitas atribut apakah ada anggota lanjut usia (usia lebih dari 65 tahun) dalam rumah tangga.	49
Tabel 4.47 Tabel Probabilitas atribut apakah ada anggota disabilitas/penyandang cacat dalam rumah tangga.	49
Tabel 4.48 Tabel Probabilitas apakah ada anggota yang	49
Tabel 4.49 Tabel Probabilitas atribut Kepala Rumah Tangga mempunyai pekerjaan tidak tetap/PHK.	49
Tabel 4.50 Tabel Probabilitas atribut Kepala rumah tangga memiliki penyakit kronis yang menyulitkan untuk bekerja.	50
Tabel 4.51 Tabel Probabilitas atribut salah Satu/Kedua Orang Tua Kandung Meninggal Dunia.....	50
Tabel 4.52 Tabel Probabilitas atribut Ada anak umur 7-18 tahun yang tidak sekolah karena alasan ekonomi.	50
Tabel 4.53 Tabel Probabilitas atribut Ayah/ibu pegawai/karyawan swasta/wiraswatawan berpenghasilan cukup.	50
Tabel 4.54 Tabel Probabilitas atribut Ayah/ibu pegawai/karyawan swasta/wiraswatawan berpenghasilan cukup (kategori penghasilan).	50
Tabel 4.55 Tabel Probabilitas atribut jumlah kepala keluarga dalam rumah tangga.	51
Tabel 4.56 Tabel Probabilitas atribut Status kepemilikan Tempat Tinggal.	51
Tabel 4.57 Tabel Probabilitas atribut luas lantai bangunan tempat tinggal kurang dari 8 m ²	51
Tabel 4.58 Tabel Probabilitas atribut bahan bangunan utama atap rumah terluas.....	51

Tabel 4.59 Tabel Probabilitas atribut bahan bangunan utama dinding rumah terluas.	52
Tabel 4.60 Tabel Probabilitas atribut bahan bangunan utama lantai rumah terluas.....	52
Tabel 4.61 Tabel Probabilitas atribut penggunaan fasilitas tempat buang air besar.	52
Tabel 4.62 Tabel Probabilitas atribut sumber penerangan utama rumah tangga.	53
Tabel 4.63 Tabel Probabilitas atribut sumber penerangan utama rumah tangga (kategori daya listrik rumah).	53
Tabel 4.64 Tabel Probabilitas atribut Sumber air minum berasal dari sumur.	53
Tabel 4.65 Tabel Probabilitas atribut bahan bakar untuk memasak.	53
Tabel 4.66 Tabel Probabilitas atribut apakah terdapat perubahan kondisi sosial ekonomi rumah tangga dari tahun sebelumnya.....	54
Tabel 4.67 Tabel Probabilitas atribut apakah rumah tangga memiliki mobil.	54
Tabel 4.68 Tabel Probabilitas atribut memiliki kapal motor.	54
Tabel 4.69 Tabel Probabilitas atribut memiliki perahu motor lebih dari atau sama dengan dua.....	54
Tabel 4.70 Tabel Probabilitas atribut memiliki sepeda motor lebih dari dua.	54
Tabel 4.71 Tabel Probabilitas atribut memiliki telepon rumah (bukan hp).	54
Tabel 4.72 Tabel Probabilitas atribut apakah rumah tangga memiliki AC.	55
Tabel 4.73 Tabel Probabilitas atribut apakah rumah tangga memiliki tabung gas lebih dari 5,5 kg.....	55
Tabel 4.74 Tabel Probabilitas atribut Anak memiliki HP berharga lebih Rp1.000.000....	55
Tabel 4.75 Tabel Probabilitas atribut rumah mewah.	55
Tabel 4.76 Tabel Probabilitas atribut usaha rumah kontrakan atau kos.	55
Tabel 4.78 Tabel Confusion Matrix Data Sampel	65
Tabel 4. 79 Tabel Confusion Matrix	66
Tabel 4. 80 Tabel Iterasi dari Seleksi Fitur <i>Backward Elimination</i>	67