

**PENERAPAN METODE *K-NEAREST NEIGHBOR*
UNTUK PREDIKSI PENERIMAAN
SISWA BARU**

(Studi Kasus : SMP N 4 Kota Gorontalo)

Oleh

NURMEYLAN H.PAKAYA

T3120031

SKRIPSI

Untuk memenuhi salah satu syarat ujian

Guna memperoleh gelar Sarjana



**PROGRAM SARJANA
TEKNIK INFORMATIKA
UNIVERSITAS ICHSAN GORONTALO
GORONTALO**

2024

PERSETUJUAN SKRIPSI

PENERAPAN METODE *K-NEAREST NEIGHBOR* UNTUK PREDIKSI PENERIMAAN SISWA BARU

(Studi Kasus : SMP Negeri 4 Kota Gorontalo)

Oleh :

NURMEYLAN H. PAKAYA

NIM : T3120031

SKRIPSI

Untuk memenuhi salah satu syarat ujian guna
memperoleh gelar Sarjan
Program Studi Teknik Informatika. Ini
telah disetujui oleh Tim Pembimbing

Pembimbing I



Rezqiwati Ishak, M.Kom
NIDN: 0903087901

Pembimbing II



Rofiq Harun, M.Kom
NIDN : 0919048404

PENGESAHAN SKRIPSI

PENERAPAN METODE *K-NEAREST NEIGHBOR* UNTUK PREDIKSI PENERIMAAN SISWA BARU

(Studi Kasus : SMP Negeri 4 Kota Gorontalo)

Oleh :

NURMEYLAN H. PAKAYA

NIM : T3120031

Diperiksa Oleh Panitia Strata(S1)

Universitas Ichsan Gorontalo

1. Ketua Penguji 1

Haditsah Annur, M.Kom

2. Anggota

Sudirman S. Panna, M.Kom

3. Anggota

Mohamad Efendi Lasulika, M.Kom

4. Anggota

Rezqiwati Ishak, M.Kom

5. Anggota

Rofiq Harun, M.Kom



Mengetahui

Dekan Fakultas Ilmu Komputer

Irvan Abraham Salihi, M.Kom
NIDN. 0928028101

Ketua Program Studi

Sudirman S. Panna, M.Kom
NIDN. 092403825

PERNYATAAN SKRIPSI

Dengan ini saya menyatakan bahwa :

1. Skripsi ini adalah asli dan belum pernah diajukan untuk mendapatkan gelar akademik (sarjana) baik di universitas Ichsan Gorontalo maupun perguruan tinggi lainnya
2. Skripsi ini adalah murni gagasan, rumusan dan penelitian saya sendiri, tanpa bantuan pihak lain, kecuali arahan Tim pembimbing
3. Dalam skripsi ini tidak terdapat karya yang telah dipublikasikan orang lain, kecuali secara tertulis dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan dicantumkan dalam daftar Pustaka
4. Pernyataan ini saya buat dengan sesungguhnya dan apa bila kemudian hari terdapat Penyimpangan dan ketidak benaran dalam pernyataan ini, maka saya bersedia menerima sanksi akademik berupa pencabutan gelar yang telah diperoleh karena Skripsi ini, serta sanksi lainnya sesuai dengan norma yang berlaku diperguruan tinggi ini.



ABSTRACT

NURMEYLAN H. PAKAYA. T3120031. THE APPLICATION OF THE K- NEAREST NEIGHBOR METHOD TO PREDICT NEW STUDENT ADMISSIONS

New student admission is one of the important aspects of school management that requires an accurate prediction method to support admission decisions. This research aims to apply the K-Nearest Neighbor (KNN) method in predicting new student admissions. The KNN method can classify data by utilizing information from K-Nearest Neighbors. The research process involves collecting previously accepted new student admission data, which includes attributes such as gender, father's occupation, and mother's occupation. The data are then divided into Training data and Testing data. The value of $K = 3$ obtains the smallest value, namely, Data 1 = 1.7916, Data 2 = 3.6105, and Data 3 = 7.7025. Through experiments with various values of K , the best K value is $K = 7$. With this K value, the - NN model has a prediction accuracy rate of 98%. These results show that the K-NN method can be effectively applied in the prediction of the new student admissions process.

Keywords: new student admission, K-Nearest Neighbor, Confusion Matrix, python

ABSTRAK

NURMEYLAN H. PAKAYA. T3120031. PENERAPAN METODE K-NEAREST NEIGHBOR UNTUK MEMPREDIKSI PENERIMAAN SISWA BARU

Penerimaan siswa baru merupakan salah satu aspek penting dalam manajemen sekolah yang memerlukan metode prediksi yang akurat untuk mendukung keputusan penerimaan. Penelitian ini bertujuan untuk menerapkan metode *K-Nearest Neighbor (KNN)* dalam prediksi penerimaan siswa baru. Metode *KNN* dipilih karena kemampuannya dalam mengklasifikasikan data dengan memanfaatkan informasi dari k tetangga terdekat. Proses penelitian melibatkan pengumpulan data penerimaan siswa baru yang telah diterima sebelumnya, yang meliputi atribut-atribut seperti jenis kelamin, pekerjaan ayah, dan pekerjaan ibu. Data tersebut kemudian dibagi menjadi data Training dan data Testing. Maka nilai $k = 3$ didapatkan nilai terkecil yaitu data ke 1 = 1,7916 data ke 2 = 3,6105 dan data ke 3 = 7,7025 Melalui eksperimen dengan berbagai nilai k , diperoleh hasil bahwa nilai k terbaik adalah $k = 7$. Dengan nilai k ini, model *KNN* mampu mencapai tingkat akurasi prediksi sebesar 98%. Hasil ini menunjukkan bahwa metode *KNN* dapat diterapkan secara efektif dalam proses prediksi penerimaan siswa baru.

Kata kunci : penerimaan siswa baru, prediksi *K-Nearest Neighbor*, akurasi *Confusion matrix* dan python



KATA PENGANTAR

Puji dan Syukur Kehadirat Allah SWT, karena berkat Rahmat dan karunianyalah penulis dapat menyelesaikan skripsi ini dengan judul, **“Penerapan metode *K-Nearest Neighbor* untuk prediksi penerimaan siswa baru”** (Studi Kasus : SMP N 4 Kota Gorontalo), sebagai salah satu syarat Ujian Akhir guna memperoleh gelar Sarjana Komputer pada Program Studi Teknik Informatika Fakultas Ilmu Komputer Universitas Ichsan Gorontalo.

Penulis menyadari sepenuhnya bahwa skripsi ini tidak mungkin terwujud tanpa bantuan dan dorongan dari berbagai pihak, baik bantuan moril maupun materil. Untuk itu, dengan segala keihlasan dan kerendahan hati, penulis mengucapkan banyak terima kasih dan penghargaan yang setinggi-tingginya kepada :

1. Ibu Dr. Dra Juriko Abdusamad, M.Si, selaku Ketua Yayasan Pengembangan Ilmu Pengetahuan Dan Teknologi (YPIPT) Universitas Ichsan Gorontalo
2. Bapak Dr. Abdu Gaffar La Tjokke, M.Si, selaku Rektor Universitas Ichsan Gorontalo
3. Bapak Irvan Abraham Salihi, M.Kom, selaku Dekan Fakultas Ilmu komputer Universitas Ichsan Gorontalo
4. Bapak Sudirman Melangi, M.Kom, selaku Wakil Dekan I Bidang Akademik dan Kemahasiswaan Fakultas Ilmu komputer Universitas Ichsan Gorontalo
5. Ibu Irma Surya Kumala Idris, M.Kom, selaku Wakil Dekan II Bidang Administrasi Umum dan Keuangan Fakultas Ilmu komputer Universitas Ichsan Gorontalo
6. Bapak Sudirman S. Panna, M.Kom, selaku Ketua Jurusan Teknik Informatika Fakultas Ilmu komputer Universitas Ichsan Gorontalo
7. Ibu Rezqiwati Ishak, M.Kom, Sebagai PembimbingUtama dalam penelitian ini yang telah membimbing penulis selama penelitian.

8. Bapak Rofiq Harun, M.Kom, sebagai Pembimbing Pendamping dalam penelitian ini yang telah membimbing penulis selama penelitian
9. Bapak dan Ibu Dosen yang telah mendidik dan membimbing serta mengajarkan berbagai disiplin ilmu kepada penulis
10. Teristimewa Ucapan Terimah Kasih kepada Kedua orang tua (Alm. Papa) mama , Kaka Perempuan dan Adik laki-laki tercinta beserta keluarga yang selalu memberikan dorongan maupun material dari awal sampai akhir perkuliahan
11. Kepada Sahabat (Inipurnawaty Rauf dan Marchella V.a.c Walone) Terimah kasih sudah selalu kebersamaan selama perkuliahan dan Terimah kasih atas suka maupun duka dalam menyelesaikan skripsi ini.
12. Kepada Teman-teman yang selalu membantu dan memberikan semangat, Teman-teman Angkatan 2020 dari jurusan Teknik Informatika, serta dari semua pihak yang ikut membantu penulis dalam menyelesaikan Skripsi ini.

Dengan demikian, penulis menyadari masih banyak kekurangan dalam penyusunan Skripsi ini, Oleh karena itu, diharapkan saran dan kritik untuk penyempurnaan penulis lebih lanjut. Semoga Skripsi ini dapat bermanfaat bagi kita semua, Aamin.

Gorontalo.....2024

Penulis

DAFTAR ISI

PERSETUJUAN	
SKRIPSI	Error! Bookmark not defined.
PENGESAHAN SKRIPSI	iii
PERNYATAAN SKRIPSI	iv
ABSTRACT	vi
ABSTRAK	vii
KATA PENGANTAR	viii
DAFTAR ISI	ix
DAFTAR GAMBAR	xiii
DAFTAR TABEL	xiii
DAFTAR LAMPIRAN	xiv
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Identifikasi Masalah	3
1.3 Rumusan Masalah	4
1.4 Tujuan Penelitian.....	4
1.5 Manfaat Penelitian.....	4
1.5.1 Manfaat Teoritis	4
1.5.2 Manfaat Praktis	4
BAB II LANDASAN TEORI	5
2.1 Tinjauan Studi	5
2.2 Tinjauan Pustaka	6
2.2.1 Penerimaan Siswa Baru.....	6
2.2.2 Prediksi	7
2.2.3 Data Mining	8
2.2.4 Algoritma <i>K-Nearest Neighbor</i>	10
2.2.4.1 Contoh Prediksi dengan Metode <i>K-Nearest Neighbor</i>	11
2.3 Pengujian Model	13
2.3.1 <i>Confusion Matrix</i>	13
2.3.2 Contoh Penerapan <i>Confusion Matrix</i>	14

2.4 Perangkat lunak Pendukung	15
2.5 Kerangka Pikir	16
BAB III METODE PENELITIAN	17
3.1 Jenis, Metode, Subjek, Waktu dan Lokasi Penelitian	17
3.2. Pengumpulan Data	17
3.3 Pemodelan	18
3.3.1 Pengembangan Model	18
3.2.3 Evaluasi Model	18
3.2.3 Tahap Implementasi	18
BAB IV HASIL PENELITIAN	19
4.1 Hasil Pengumpulan Data	19
4.2 Hasil Pemodelan	20
4.2.1 Data Preprocessing	20
4.2.2 Pembagian Data Training dan Data Testing	22
4.2.3 Pemodelan dengan <i>K-Nearest Neighbor</i>	22
4.3. Pengujian Model Dan Evaluasi Kinerja Model	23
4.3.1 Implementasi Hasil Pemodelan	24
4.3.2 Implementasi Data Processing	24
4.3.3 Implementasi Pembagian Data Training dan Data Testing	32
4.3.4 Implementasi Pemodelan	32
4.3.5 Implementasi Pengujian Model dan Evaluasi Kinerja Model	33
4.3.6 Implementasi Hasil Prediksi	35
BAB V PEMBAHASAN PENELITIAN	37
5.1 Pembahasan Kinerja Model	37
5.2 Pembahasan Hasil Pemodelan	40
BAB VI KESIMPULAN DAN SARAN	44
6.1 Kesimpulan	44
6.2 Saran	44
DAFTAR PUSTAKA	45

DAFTAR LAMPIRAN

LAMPIRAN 1 PROGRAM

LAMPIRAN 2 SURAT TELAH MELAKUKAN PENELITIAN

LAMPIRAN 3 SURAT KETERANGAN BEBAS PUSTAKA

LAMPIRAN 4 SURAT REKOMENDASI BEBAS PLASGIASI

LAMPIRAN 5 HASIL TURNITIN

LAMPIRAN 6 DAFTAR RIWAYAT HIDUP

DAFTAR GAMBAR

Gambar 2.1 Proses <i>Knowledge Discoveryin Database</i>	8
Gambar 2.2 Menunjukkan hubungan data mining dengan area-area lain	9
Gambar 5.1 <i>Confusion Matrix</i> model <i>KNN</i> untuk $k=3$	37
Gambar 5.2 <i>Confusion Matrix</i> model <i>KNN</i> untuk $k=5$	38
Gambar 5.3 <i>Confusion Matrix</i> model <i>KNN</i> untuk $k=7$	39
Gambar 5.4 Hasil Penyesuaian Atribut	40
Gambar 5.5 Pembagian Data Training dan Data Testing	41

DAFTAR TABEL

Tabel 1.1	Jumlah Penerimaan Siswa Baru	2
Tabel 2.1	Penelitian Terkait.....	5
Tabel 2.2	Atribut Identitas Siswa.....	7
Tabel 2.3	Pembobotan Atribut.....	11
Tabel 2.4	Sampel Data Training	12
Tabel 2.5	Sampel Data Testing.....	12
Tabel 2.6	<i>Confusion Matrix</i>	13
Tabel 2.7	Contoh Kasus <i>Confusion Matrix</i>	14
Tabel 4.1	Dataset Awal Penerimaan Siswa Baru.....	19
Tabel 4.2	Konversi Atribut jenis Kelamin.....	20
Tabel 4.3	Konversi Atribut Pekerjaan Ayah	20
Tabel 4.4	Konversi Atribut Pekerjaan Ibu	21
Tabel 4.5	Konversi Atribut jenis Awal Penerimaan siswa baru	21
Tabel 4.6	Pembagian Data Training dan Data Testing	22
Tabel 4.7	Sampul Perhitungan Jarak Encludian	22
Tabel 5.1	<i>Classification Report</i> Model <i>KNN</i> untuk $k=3$	37
Tabel 5.2	<i>Classification Report</i> Model <i>KNN</i> untuk $k=5$	37
Tabel 5.3	<i>Classification Report</i> Model <i>KNN</i> untuk $k=7$	39
Tabel 5.4	Pengujian Data Testing Baru.....	41
Tabel 5.5	Hasil Pengujian Data Testing Baru.....	43

DAFTAR LAMPIRAN

Lampiran 1 Program

Lampiran 2 Surat Telah Melakukan Penelitian

Lampiran 3 Surat Keterangan Bebas Pustaka

Lampiran 4 Surat Rekomendasi Bebas Plagiasi

Lampiran 5 Hasil Turnitin

Lampiran 6 Daftar Riwayat Hidup

BAB I

PENDAHULUAN

1.1 Latar Belakang

Pendidikan dan sekolah adalah dua istilah yang berbeda. Pendidikan merupakan sistem dan proses dalam mendidik, baik secara formal, informal, maupun nonformal. Sementara itu, sekolah merupakan tempat melaksanakan pembelajaran atau Pendidikan secara formal. Sekolah adalah sebuah Lembaga yang dirancang untuk pengajaran siswa atau murid dibawah pengawasan guru. Sekolah memiliki peran yang sangat vital dalam mengembangkan potensi individu manusia. Selain itu, institusi pendidikan ini juga bertanggung jawab untuk menyiapkan individu agar siap menghadapi masa depan mereka dengan lebih baik, baik dalam hal kesejahteraan pribadi maupun kontribusi mereka dalam masyarakat, bangsa, dan negara. Sekolah adalah suatu entitas atau organisasi di mana proses pembelajaran dan pengajaran berlangsung, serta tempat dimana pengetahuan diberikan dan diterima.[1]

Penerimaan siswa baru (PSB) pada umumnya dapat diartikan sebagai suatu proses administrasi yang terjadi setiap tahun untuk seleksi calon siswa berdasarkan nilai akademik agar dapat melanjutkan pendidikan pada jenjang yang lebih tinggi.[2] Kegiatan pendaftaran siswa baru merupakan kegiatan rutin yang dilakukan di sekolah pada setiap ajaran baru. Penerimaan siswa baru merupakan suatu agenda rutinitas yang dilakukan oleh Lembaga Pendidikan nasional disemua jenjang. Penerimaan siswa baru adalah suatu tahap yang rutin terjadi setiap tahun sebelum dimulainya tahun pelajaran baru, yang dilaksanakan oleh setiap jenis sekolah, baik yang dikelola oleh pemerintah maupun swasta.

SMP N 4 Kota Gorontalo yakni Masyarakat kecamatan kota barat Kota Gorontalo telah lama menginginkan sebuah sekolah menengah tingkat pertama. Tamatan sekolah Dasar dari kecamatan ini selaku merasa kecewa karena Sebagian besar tidak ter terima di SMP yang jauhnya kurang lebih 5 km dari tempat tinggal lulusannya. Antara tahun 1962/1963 siswa lulusan SD berasal dari kecamatan kota barat masuk (mendaftar) di SMP Negeri 2 Kota Gorontalo, di daftarkan tersendiri,

dijadikan satu kelas dan diberi nama SMP N 2 kelas jauh potanga. Tahun ajaran 1963/1964 kelas jauh potanga ini telah memiliki kelas I dan kelas II, siswa menggunakan aula SDN No.2 Potanga Kecamatan Kota Barat. Tahun ajaran 1965/1966 masyarakat dan orang tua siswa atau himbauan pemerintah setempat, berhasil mendirikan sebuah gedung sekolah yang dapat menampung satu ruangan kelas I, satu ruangan kelas II, satu ruangan kelas III gedung yang dibangun tersebut berlokasi dikelurahan kota Barat kecamatan Kota Barat Kota Gorontalo.

Berikut ini data jumlah penerimaan siswa baru di SMP N 4 Kota Gorontalo dari tahun 2012-2023.

Tabel 1.1 jumlah penerimaan siswa baru

No	Tahun	Jumlah Penerimaan Siswa baru
1	2012	320
2	2013	254
3	2014	236
4	2015	231
5	2016	209
6	2017	340
7	2018	243
8	2019	235
9	2020	233
10	2021	238
11	2022	218
12	2023	224

Sumber : (SMP N 4 Kota Gorontalo, 2023)

Berdasarkan pada tabel di atas pendaftaran penerimaan siswa baru di lihat dari jarak rumah ke sekolah. Karena adanya aturan yang ditetapkan, siswa yang mendaftar akan diseleksi yang diterima atau tidak disekolah tersebut. Dari permasalahan tersebut maka diperlukan sistem untuk pengambilan keputusan

dalam hal memprediksi penerimaan siswa baru SMP N 4 Kota Gorontalo berikutnya.

Metode *K-Nearest Neighbour (KNN)* merupakan algoritma yang digunakan untuk mengklasifikasi terhadap object berdasarkan jarak eculidien paling dekat dengan object tersebut. Algoritma *K-Nearest Neighbor (KNN)* adalah sebuah metode untuk melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut.[3]

Berdasarkan penelitian yang dilakukan oleh Adhitya Rahmat D.N tentang “Implementasi metode *K-Nearest Neighbor* untuk seleksi calon karyawan baru” Hasil pengujian perhitungan algoritma *K-Nearest Neighbor* dengan nilai $K=7$ didapat nilai akurasi sebesar 91%, nilai presisi sebesar 87% dan recall 100%

Berdasarkan hasil dari penelitian terkait diatas, maka penggunaan metode *K-Nearest Neighbor* ini sangat cocok digunakan pada penelitian ini dikarenakan Algoritma ini bisa memberikan nilai akurasi yang baik untuk melakukan proses prediksi penerimaan siswa baru

Adanya permasalahan diatas, maka penulis tertarik untuk melakukan penelitian yang diberi judul **“Penerapan Metode *K-Nearest Neighbor(K-NN)* untuk prediksi penerimaan siswa baru (Studi Kasus : SMP N 4 Kota Gorontalo).**

1.2 Identifikasi Masalah

Berdasarkan latar belakang diatas, maka identifikasi masalahnya adalah :

1. Pendaftaran penerimaan siswa baru dilihat dari jarak rumah ke sekolah
2. Belum adanya suatu sistem prediksi yang digunakan oleh pihak sekolah SMP N 4 Kota Gorontalo dalam memprediksi penerimaan siswa baru

1.3 Rumusan Masalah

Dari uraian latar belakang masalah tersebut, dapat dirumuskan masalah pokok yang berkaitan yaitu:

1. Bagaimana penerapan metode *K-Nearest Neighbor* dalam memprediksi penerimaan siswa baru
2. Seberapa akurat metode *K-Nearest Neighbor* dalam memprediksi Penerimaan siswa baru?

1.4 Tujuan Penelitian

Adapun tujuan dari penelitian ini:

1. Menerapkan metode *K-Nearest Neighbor* dalam memprediksi penerimaan siswa baru
2. Mengevaluasi akurasi metode *K-Nearest Neighbor* dalam memprediksi penerimaan siswa baru

1.5 Manfaat Penelitian

1.5.1 Manfaat Teoritis

Penelitian ini diharapkan dapat memberikan kontribusi pemikiran, pengetahuan dan wawasan keilmuan dibidang teknologi komputer khususnya dalam penerapan metode K-nn untuk prediksi jumlah penerimaan siswa baru

1.5.2 Manfaat Praktis

Diharapkan hasil penelitian yang dibuat ini dapat memberikan masukan atau referensi serta menambah pengetahuan bagi mahasiswa pada umumnya dan khususnya bagi program Studi Fakultas Ilmu Komputer

BAB II

LANDASAN TEORI

2.1 Tinjauan Studi

Prediksi menggunakan Metode *K-Nearest Neighbor* merupakan bidang penelitian yang telah banyak dikembangkan saat ini. Berikut penelitian terkait yang menjadi referensi.

Tabel 2.1 Penelitian Tentang Prediksi dengan metode *KNN*

No	Peneliti	Judul	Tahun	Metode	Hasil
1	Kartika et. al	Penentuan siswa berprestasi menggunakan metode <i>KNN</i> dengan weighted product	2017	<i>K-Nearest Neighbor (KNN)</i>	Pengujian perangkingan pertama memiliki tingkat akurasi 66.67%. pengujian perangkingan kedua memiliki tingkat akurasi sebesar 88.89%. prngujian perangkingan ketiga memiliki tingkat akurasi sebesar 100%. [3]
2	Sumarlin	Implementasi algoritma <i>K-Nearest Neighbor</i> sebagai pendukung keputusan klasifikasi penerimaan beasiswa peningkatan	2015	<i>K-Nearest Neighbor (KNN)</i>	Hasil akurasi yang diperoleh untuk beasiswa peningkatan prestasi akademik mencapai 88.33% dengan nilai <i>Area Under Curva(AUC)</i> 0.925 dari 227 record dataset, sedangkan akurasi yang diperoleh untuk beasiswa bantuan

		prestasi akademik			belajar mahasiswa mencapai 90% dengan nilai AUC 0.937 Dari 183 record dataset, akurasi yang diperoleh untuk gabungan beasiswa PPA dan BBM mencapai 85.56% dan nilai AUC 0.958[3]
No	Peneliti	Judul	Tahun	Metode	Hasil
3	Adhitya Rahmat D.N	Implementasi metode <i>K-Nearest Neighbor</i> untuk seleksi calon karyawan baru	2020	<i>K-Nearest Neighbor (KNN)</i>	Hasil pengujian perhitungan algoritma <i>K-Nearest Neighbor</i> dengan nilai $K=7$ didapat nilai akurasi sebesar 91%, nilai presisi sebesar 87% dan recall 100%[1]

2.2 Tinjauan Pustaka

2.2.1 Penerimaan Siswa Baru

Penerimaan siswa baru (PSB) pada umumnya dapat diartikan sebagai suatu proses administrasi yang terjadi setiap tahun untuk seleksi calon siswa berdasarkan nilai akademik agar dapat melanjutkan pendidikan pada jenjang yang lebih tinggi.[2] Penerimaan siswa baru merupakan suatu agenda rutinitas yang dilakukan oleh Lembaga Pendidikan nasional di semua jenjang. Kegiatan pendaftaran siswa baru merupakan kegiatan rutin yang dilakukan di sekolah pada setiap ajaran baru. Hal ini merupakan peristiwa penting bagi suatu sekolah, karena peristiwa ini merupakan titik awal yang menentukan kelancaran tugas suatu sekolah. Penerimaan siswa baru dilakukan bukanlah hal yang ringan. Sekolah harus menyiapkan strategi-

strategi yang tepat dalam menjalankannya, supaya dapat menarik siswa-siswa yang berkualitas yang mana input sekolah juga bisa lebih baik sehingga proses belajar bisa maksimal dan kualitas sekolah meningkat.[4]

Pada penelitian ini yang menjadi atribut identitas siswa dalam melakukan prediksi sesuai dengan indikator penerimaan siswa baru berdasarkan Latar Belakang, seperti pada tabel dibawah ini :

Table 2.2 Atribut Identitas Siswa

No	Identitas Siswa
1	Nama Siswa
2	Jenis Kelamin
3	Jarak Rumah Ke Sekolah
4	Pekerjaan orang tua
5	Nilai Raport

2.2.2 Prediksi

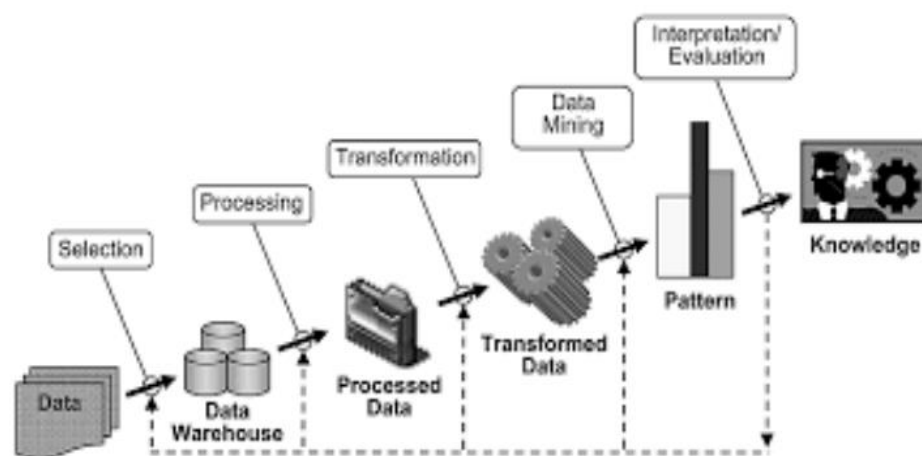
Prediksi merupakan suatu metode sistematis untuk memproyeksikan kemungkinan terjadinya peristiwa di masa depan dengan merujuk pada data masa lalu dan saat ini, dengan tujuan mengurangi kesalahan dalam perbandingan antara hasil prediksi dengan realitas yang sebenarnya.

Menurut Herdianto (2013:8) Prediksi adalah “suatu proses memperkirakan secara sistematis tentang sesuatu yang paling mungkin terjadi dimasa depan berdasarkan informasi masa lalu dan sekarang yang dimiliki, agar kesalahannya (selisih antara sesuatu yang terjadi dengan hasil perkiraan) dapat diperkecil. Prediksi tidak harus memberikan jawaban secara pasti yang akan terjadi, melainkan berusaha untuk mencari jawaban sedekat mungkin yang akan terjadi”. [5] Prediksi tidak bertujuan memberikan jawaban pasti tentang apa yang akan terjadi, tetapi berusaha untuk mencari perkiraan yang paling mendekati kemungkinan yang akan terjadi.

2.2.3 Data Mining

Data mining merupakan suatu proses yang memanfaatkan teknik statistik, matematika, kecerdasan buatan dan machine learning guna mengekstraksi serta mengidentifikasi informasi berharga serta pengetahuan terkait dari berbagai basis data yang berukuran besar. [3] Data Mining adalah serangkaian proses untuk menggali nilai tambah dari suatu Kumpulan data berupa pengetahuan yang selama ini. Data mining merupakan suatu Langkah dalam *Knowledge Discovery in database*(KDD).

Knowledge Discovery in Database(KDD) adalah tahapan global dari penggalian data yang terdiri dari beberapa langkah / tahapan, diantaranya pemilihan data mentah, pra pengolahan data (preprocessing), transformasi data, data mining, dan evaluasi hasil dari pengolahan informasi. Maka dapat dikatakan bahwa data mining merupakan bagian dari *Knowledge Discovery in Database* (KDD). [6]



Gambar 2.1 Proses *Knowledge Discovery in Database*(KDD)

Sumber : Amalia, 2018

Perkembangan ini, data mining semakin banyak digunakan untuk mengatasi permasalahan yang berkaitan dengan bidang Pendidikan. Data mining digunakan untuk menggali data yang telah ada dengan tujuan membangun sebuah model, lalu memanfaatkan model tersebut untuk mengidentifikasi pola data lain yang tidak

terdapat dalam basis data yang tersimpan. Salah satu teknik dalam data mining adalah teknik klasifikasi. Klasifikasi adalah metode pembelajaran yang digunakan untuk memprediksi nilai dari variabel target yang termasuk dalam kategori tertentu.

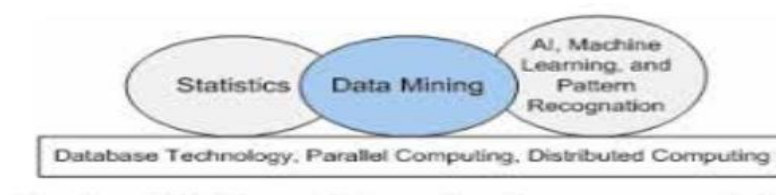
Kegunaan dari data mining adalah untuk mengidentifikasi dan mengklasifikasikan pola yang perlu ditemukan dalam data. . Secara umum semua operasi dalam data mining dapat dikelompokkan menjadi dua metode yaitu metode deskriptif dan metode prediktif.[3]

Adapun operasi-operasi dan Teknik-teknik yang berhubungan :

1. Operasi *Predictive Modeling* : (*classification, value prediction*)
2. *Database segmentation* : (*demographic claustring,neural claustring*)
3. *Link analysis* : (*association discovery, sequential pattern descovey, similar time sequencediscovey*)
4. *Deviation detection* : (*statistic, visualization*)

Hasil dari data mining sering digabungkan dengan sistem pendukung keputusan (Decision Support System, DSS). Sebagai contoh, dalam aplikasi bisnis, informasi yang diperoleh dari data mining dapat diintegrasikan dengan alat manajemen kampanye produk untuk melaksanakan promosi pemasaran yang lebih efektif dan dapat diuji. Integrasi semacam ini melibatkan langkah-langkah pasca-pengolahan yang memastikan hanya hasil yang sah dan bermanfaat yang akan dimasukkan ke dalam DSS. Salah satu tugas pasca-pengolahan adalah visualisasi data, yang memungkinkan analisis untuk menjelajahi data dan hasil data mining dari berbagai sudut pandang. Selama proses pasca-pengolahan, ukuran statistik dan metode pengujian hipotesis dapat digunakan untuk mengidentifikasi dan menghapus hasil data mining yang tidak valid.

Gambar 2.2 menunjukkan hubungan data mining dengan area-area lain



Sumber : (Tan et al, 2005)

Dalam konteks ini, data mining secara spesifik mengadopsi konsep-konsep seperti (1) pengambilan sampel, estimasi, dan pengujian hipotesis dari domain statistik, dan (2) algoritma pencarian, teknik pemodelan, dan teori pembelajaran dari bidang kecerdasan buatan, pengenalan pola, dan pembelajaran mesin.

Data mining juga mengintegrasikan konsep-konsep dari bidang lain, termasuk optimisasi, komputasi evolusioner, teori informasi, pemrosesan sinyal, visualisasi, dan temu kembali informasi. Beberapa bidang lain juga memberikan kontribusi penting dalam data mining, seperti sistem basis data yang diperlukan untuk menyediakan tempat penyimpanan yang efisien, indeks, dan pemrosesan query.

2.2.4 Algoritma *K-Nearest Neighbor*

Metode *K-Nearest Neighbor* adalah salah satu Teknik pengklasifikasian data yang memanfaatkan perhitungan tingkat kedekatan antar data baru dan data yang telah ada berdasarkan penyesuaian bobot. Pendekatan ini terbukti lebih efisien dalam memisahkan kategori siswa. Metode *K-Nearest Neighbor* sangat Tangguh dan efektif apabila training datanya besar. Secara umum *K-Nearest Neighbor* digunakan untuk mengelompokkan suatu variabel dengan kategori[7].

Dibawah ini merupakan Langkah-langkah pada metode K-Nearest Neighbor :

- a) Menentukan nilai 'K' untuk menentukan banyaknya jumlah tetangga dengan data uji
- b) Kemudian menghitung jarak antara data uji terhadap setiap data latih dengan *Enculidean Distance* (menghitung nilai *Enculidean Distance*)
- c) Kemudian ukuran hasil perhitungan jarak *Enculidean Distance*
- d) Setelah itu menghitung frekuensi klasifikasi terbanyak dari data yang sudah dihitung dan diurutkan serta menentukan klasifikasi yang frekuensinya paling banyak
- e) Yang terakhir adalah memiliki klasifikasi dengan frekuensi tertinggi untuk menentukan klasifikasi terhadap data uji.

Persamaan dibawah adalah rumus untuk menghitung jarak *eculidean distance* pada metode *K-NN*

$$d(P, Q) = \sqrt{\sum_i^n (P_i - Q_i)^2}$$

Keterangan

$d(P, Q)$: Jarak eculidean
 n : Jumlah data latih
 p : Inputan data ke-1 dari data uji
 Q : Inputan data ke-1 dari data latih

eculidean distance adalah salah satu Teknik yang digunakan dalam metode klasifikasi untuk mengukur kesamaan antara data uji dengan data lain dalam data set.[1]

2.2.4.1 Contoh Prediksi dengan Metode *K-Nearest Neighbor*

Berikut ini merupakan contoh penerapan metode *KNN* dengan studi kasus sebagai berikut :

Contoh : Prediksi Siswa penerima BSM (Bantuan Siswa Miskin)

Contoh data siswa yang menerima BSM maupun yang tidak menerima BSM. Setiap atribut memiliki bobot nilai, seperti terlihat pada Tabel dibawah ini:

Tabel 2.3 Pembobotan Atribut

No	Atribut	Bobot
1	Tanggungan(b)	0,75
2	Pekerjaan Orangtua(d)	1
3	Penghasilan Orangtua	0,5

Tabel 2.4 Sampel Data Training

Siswa	Tanggapan Orang Tua	Pekerjaan Orang Tua	Penghasilan Orang tua	Status BSM
Amin	3	Swasta	<1.500.000	Dapat
Bela	2	Petani	<750.000	Dapat
Cindy	1	Wiraswasta	<1.500.000	Tidak
Desi	2	PNS	>2.500.000	Tidak

Misalkan ada kasus baru pada data testing dengan nilai Atribut seperti pada table 2.3 kasus baru tersebut akan dihitung kedekatan dengan kasus lama yang terdapat pada data training table 2.4.

Tabel 2.5 Sampel Data Testing

Siswa	Tanggapan Orang Tua	Penghasilan Orang tua	Pekerjaan Orang Tua
Aldi	1	<750.000	Buruh

Perhitungan kedekatan kasus baru pada data testing(Tabel 2.3) dengan 4 kasus lama pada data training (Tabel 2.2) yaitu:

$$\begin{aligned}
 \text{Jarak Siswa Amin} &= \frac{(axb)+(cxd)+(exf)}{b+d+f} \\
 &= \frac{(0,2 \times 0,75) + (0 \times 1) + (0,6 \times 0,5)}{0,75 + 1 + 0,5} \\
 &= \frac{0,15 + 0 + 0,3}{2,25} \\
 &= 0,2
 \end{aligned}$$

$$\begin{aligned}
 \text{Jarak Siswa Bela} &= \frac{(axb)+(cxd)+(exf)}{b+d+f} \\
 &= \frac{(0,6 \times 0,75) + (0,8 \times 1) + (1 \times 0,5)}{0,75 + 1 + 0,5} \\
 &= \frac{0,45 + 0,8 + 0,5}{2,25} \\
 &= 1,75
 \end{aligned}$$

$$\begin{aligned}
 \text{Jarak Siswa Cindy} &= \frac{(axb)+(cxd)+(exf)}{b+d+f} \\
 &= \frac{(1 \times 0,75) + (0,2 \times 1) + (0,6 \times 0,5)}{0,75 + 1 + 0,5} \\
 &= \frac{0,75 + 0,2 + 0,3}{2,25} \\
 &= 0,55
 \end{aligned}$$

$$\begin{aligned}
 \text{Jarak Siswa Desi} &= \frac{(axb)+(cxd)+(exf)}{b+d+f} \\
 &= \frac{(0,6 \times 0,75) + (0,1 \times 1) + (0 \times 0,5)}{0,75 + 1 + 0,5} \\
 &= \frac{0,45 + 0 + 0}{2,25} \\
 &= 0,2
 \end{aligned}$$

Jadi Sample data testing siswa Bernama Aldi akan mendapat BSM karena pendekatan ke siswa B dengan jarak 1,75

2.3 Pengujian Model

2.3.1 Confusion Matrix

Confusion Matrix merupakan algoritma yang bertujuan untuk menganalisis seberapa baik metode klasifikasi yang digunakan dalam mengklasifikasi. *Confusion matrix* memberikan keputusan yang diperoleh dalam training dan testing. *Confusion matrix* memberikan penilaian dalam performance klasifikasi berdasarkan objek dengan benar atau salah. Tingkat akurasi dapat dihitung dengan melakukan persentase jumlah klasifikasi yang benar di bagi dengan jumlah data testing yang dijadikan sebagai data uji pada klasifikasi.[1] Terdapat beberapa istilah pada Algoritma Confusion Matrix yaitu TP dan TN yang berarti hasil klasifikasi terdeteksi benar, sedangkan FP dan FN yang berarti hasil klasifikasi terdeteksi salah

Tabel 2.6 Confusion Matrix

		Ya	Tidak
Predictive	Ya	TP	FN
Class	Tidak	FP	TN
Total		P'	N'

Keterangan:

TP (True Positif) : Data kelas positif yang diklasifikasikan sebagai kelas positif.

TN (True Negatif) : Data kelas negatif yang diklasifikasikan sebagai kelas negatif.

FP (False Positif) : Data kelas positif yang diklasifikasikan sebagai kelas negatif.

FN (False Negatif) : Data kelas negatif yang diklasifikasikan sebagai kelas positif.

2.3.2 Contoh Penerapan *Confusion Matrix*

Confusion matrix menghitung test record yang salah dan benar dari prediksi-prediksi. Pada 3 value class, ada tiga kategori output. Contoh kasus confusion matrix, pada sistem pengenalan predicted Garpu, Sendok, dan Sutil, terdapat data set yang menghasilkan prediksi dalam confusion matrix sbb:

Tabel 2.7 Contoh Kasus Confusion Matrix

		Actual		
		Sendok	garpu	Sutil
Predicted	Sendok	4	1	1
	Garpu	2	5	2
	Sutil	1	2	10

Pada *Confusion Matrix* di atas, Actual sebenarnya ada 7 sendok, namun sistem memprediksi ada 4 sendok, 2 garpu, dan 1 sutil

Rumus yang digunakan :

$$Accuracy = \frac{\text{Total correct predictions}}{\text{Total correct and incorrect prediction}}$$

$$= \frac{f_{11} + f_{22} + f_{33}}{F_{11} + F_{12} + F_{13} + F_{21} + F_{22} + F_{23} + F_{31} + F_{32} + F_{33}}$$

$$ErrorRate = \frac{\text{Total incorrect predictions}}{\text{Total correct and incorrect prediction}}$$

$$= \frac{f_{11} + f_{22} + f_{33}}{F_{11} + F_{12} + F_{13} + F_{21} + F_{22} + F_{23} + F_{31} + F_{32} + F_{33}}$$

Dari table confusion matrix di atas dapat dihitung Accuracy dan error ratenya

$$Accuracy = \frac{Total\ correct\ predictions}{Total\ correct\ and\ incorrect\ prediction}$$

$$= \frac{4 + 5 + 10}{4 + 1 + 1 + 2 + 5 + 2 + 1 + 2 + 10}$$

$$Accuracy = \frac{19}{28} = 65.55\%$$

$$ErrorRate = \frac{Total\ incorrect\ predictions}{Total\ correct\ and\ incorrect\ prediction}$$

$$= \frac{1 + 1 + 2 + 2 + 1 + 2}{4 + 1 + 1 + 2 + 5 + 2 + 1 + 2 + 10}$$

$$Error\ Rate = 9/28 = 34,4\%$$

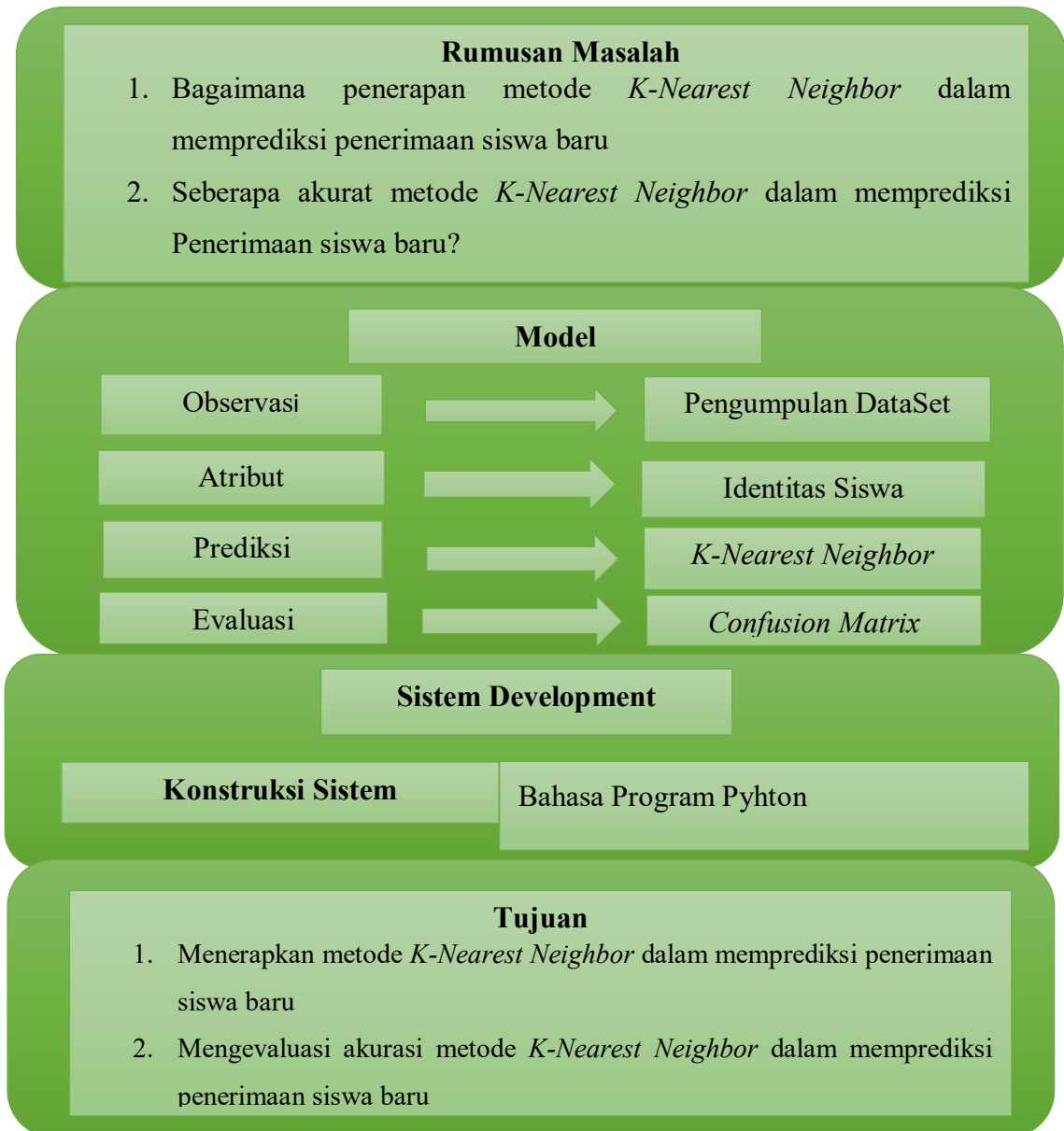
2.4 Perangkat lunak Pendukung

Perangkat lunak pendukung yang digunakan penulis dalam penelitian ini yaitu *Python*.

Python merupakan bahasa pemrograman tingkat tinggi, penulisan kode/sintaks lebih sederhana, bersifat *open-source*, *cross-platfrom* dan cocok digunakan untuk Data Scientist serta didukung banyak library seperti *Numpy*, *SciPy*, *Pandas*, *Scikit-Learn*, dan *Matplotlib*.

2.5 Kerangka Pikir

Gambar 2.3 Bagan Kerangka Pikir



BAB III

METODE PENELITIAN

3.1 Jenis, Metode, Subjek, Waktu dan Lokasi Penelitian

Berdasarkan dari tingkat penerapan maka, penelitian ini merupakan penelitian terapan. Dipandang dari jenis informasi yang diolah maka penulisan dalam hal ini yaitu bersifat kuantitatif. Sedangkan terhadap perilaku terhadap item, maka merupakan penelitian konfirmatori berdasarkan perilaku data

Tujuan penulisan ini antara lain prediksi penerimaan siswa baru, studi kasus SMP N 4 Kota Gorontalo menggunakan metode *KNN*. Penelitian ini dimulai dari Juli – September yang berlokasi pada SMP N 4 Kota Gorontalo

3.2. Pengumpulan Data

Data yang digunakan dalam penelitian ini ada 2 (dua) yaitu jenis yakni data primer dan sekunder. Data primer adalah jenis data yang dikumpulkan oleh peneliti secara langsung dari sumber utama, bisa melalui wawancara, survei dan Data Sekunder berasal dari penelitian kepustakaan

1. Data sekunder

Data sekunder merupakan data yang sudah tersedia sehingga kita tinggal mencari dan mengumpulkan. Data sekunder dari penelitian ini adalah metode kepustakaan, yaitu teori-teori yang sudah ada. Berupa teori-teori tentang Data mining, metode *K-NN* maupun tentang prediksi penerimaan siswa baru

2. Data Primer

Data primer adalah data yang dikumpulkan langsung oleh peneliti di SMP N 4 Kota Gorontalo data yang peneliti dapatkan berupa data-data yang masih secara manual.

Sedangkan cara pengumpulan data pada penelitian ini digunakan beberapa cara diantaranya :

1. Observasi

Pada tahap ini penulis melakukan pengamatan terhadap dokumen-dokumen sekolah khususnya dokumen-dokumen penerimaan siswa baru

2. Wawancara

Pada tahap ini penulis melakukan tanya jawab dengan ibu Meylan S.Kadir S.Pd selaku ketua tata usaha mengenai semua yang berkaitan dengan proses penerimaan siswa baru kepada pihak SMP N 4 Kota Gorontalo guna mendapatkan informasi yang diperlukan

3. Studi Literatur

Studi literatur yang dilakukan dalam penelitian ini adalah dengan mengumpulkan data-data yang bersumber dari berbagai buku yang menjadi referensi dan pencarian dengan media internet untuk memperoleh data-data tambahan yang diperlukan

3.3 Pemodelan

3.3.1 Pengembangan Model

Data prosedur atau langkah-langkah pokok dalam prediksi penerimaan siswa baru menggunakan metode *KNN* dengan menggunakan alat bantu tools Python untuk menguji kinerja sistemnya

3.2.3 Evaluasi Model

Evaluasi model dalam metode *K-Nearest Neighbor* bertujuan untuk menilai seberapa baik model tersebut dapat memprediksi kelas atau nilai target dari data baru dan model yang telah dihasilkan kemudian dievaluasi dengan menggunakan *Precision*, *Recall*, dan *f1-score* untuk mengetahui hasil akhir prediksi

3.2.3 Tahap Implementasi

Tahap implementasi sistem (*system implementasion*) merupakan tahap meletakkan sistem supaya siap untuk di operasikan pada masyarakat, dalam hal ini mengimplementasikan atau menerapkan prediksi penerimaan siswa baru menggunakan metode *K-Nearest Neighbor* di SMP N 4 Kota Gorontalo.

BAB IV

HASIL PENELITIAN

4.1 Hasil Pengumpulan Data

Dataset penerimaan siswa baru pada kasus ini antara lain, data yang diteliti diperoleh dari SMP N 4 Kota Gorontalo sebanyak 472 record dalam bentuk file excel. Bentuk dataset ditunjukkan pada Tabel 4.1

Tabel 4.1 Dataset Awal Penerimaan Siswa baru

No	Nama Siswa	JK	Jarak Rumah (KM)	Pekerjaan Ayah	Pekerjaan Ibu	Nilai Raport	Status
1	ABDUL RAHMAT DUNGGIO	L	2	Wiraswasta	Karyawan Swasta	86.91	Diterima
2	ALFADLI RAHMAN	L	2	Karyawan Swasta	IRT	84.83	Diterima
3	SYAHRIL PAKAYA	L	4	Buruh	IRT	84.23	Tidak Diterima
4	AQRIL PUTRA NENTO	L	2	Nelayan	IRT	87.83	Diterima
5	ARDIL TAUFIK ABDULLAH	L	3	Pedagang	IRT	85.73	Diterima
6	FAREL PONINI	L	2	Karyawan Swasta	IRT	87.02	Diterima
7	ISMAIL S HARUN	L	2	Nelayan	IRT	82	Diterima
8	MOH.ALIF FARELYANSYAH LAPAMONA	L	2	Wiraswasta	Karyawan Honorer	86.53	Diterima
9	MOH.REZKI MARDJUN	L	1	Buruh	IRT	85.46	Diterima
10	CITRA UMAR	P	5	Buruh	IRT	81.47	Tidak Diterima
11	MOH.WAHYU KURNIA S.UMAR	L	2	Wiraswasta	IRT	86.83	Diterima
12	MOH.ADIEL ULYA K TAHIR	L	1	Karyawan Swasta	Karyawan Swasta	85.73	Diterima
13	MUHAMAD FADIL HILIMI	L	1	Buruh	IRT	87.33	Diterima
14	RAMADANI ADAM	L	2	Buruh	IRT	87.54	Diterima
15	SAHRUL SOLEMAN	L	1	Wiraswasta	IRT	86.84	Diterima
16	ALFAREL	L	7	Sopir	IRT	80.56	Tidak Diterima
17	AISYIYAH LIA KOPITOY	P	2	Wiraswasta	IRT	87.92	Diterima
18	FADHILAH NUR ASYFA S.B	P	1	Buruh	IRT	86.82	Diterima
.....							
472	ADNAN POHA	L	1	PNS	PNS	86.33	Diterima

4.2 Hasil Pemodelan

Sebelum dilakukan pemodelan dengan metode *K-Nearest Neighbor* terlebih dahulu dilakukan beberapa preprocessing data terhadap dataset pada Tabel 4.1 diatas khususnya untuk atribut Dataset awal penerimaan siswa baru dilakukan konversi data dari jenis kategorikal menjadi numerik supaya datanya menjadi satu jenis type data yaitu numerik

4.2.1 Data Preprocessing

1. Konversi Data untuk Atribut Jenis Kelamin

Konversi nilai atribut untuk atribut jenis kelamin ditujukan pada Tabel 4.2 berikut :

Tabel 4.2 Konversi Atribut Jenis Kelamin

No.	Jenis Kelamin	Nilai
1	L	1
2	P	2

Berikut ini nilai konversi Atribut jenis kelamin didapatkan dalam sistem bahwa nilai laki-laki bernilai 1 dan perempuan bernilai 2

2. Konversi Data untuk Atribut Pekerjaan Ayah

Konversi nilai atribut untuk atribut Pekerjaan Ayah ditujukan pada Tabel 4.3 berikut :

Tabel 4.3 Konversi Atribut Pekerjaan Ayah

No.	Pekerjaan Ayah	Nilai
1	Buruh	1
2	Petani	2
3	Nelayan	3
4	Tukang	4
5	Sopir	5
6	Honorer	6
7	Karyawan Swasta	7
8	Pedagang	8
9	Wiraswasta	9
10	Pensiunan	10
11	TNI/Polri	11

No	Pekerjaan Ayah	Nilai
12	PNS	12
13	Pekerjaan Lainnya	13

Pada Tabel yang diatas menjelaskan bahwa nilai konversi atribut pekerjaan ayah diubah dalam bentuk numerik

3. Konversi Data untuk Atribut Pekerjaan Ibu

Konversi nilai atribut untuk atribut Pekerjaan Ibu ditujukan pada Tabel 4.4 berikut :

Tabel 4.4 Konversi Atribut Pekerjaan Ibu

No.	Pekerjaan Ibu	Nilai
1	IRT	1
2	Karyawan Honorer	2
3	Karyawan Swasta	3
4	Guru	4
5	Wiraswasta	5
6	PNS	6

Pada Tabel yang diatas menjelaskan bahwa nilai konversi atribut pekerjaan Ibu diubah dalam bentuk numerik

Setelah dilakukan konversi untuk semua atribut yang bernilai kategorikal, maka hasil akhir dataset yang akan digunakan ditunjukkan pada Tabel 4.5

Tabel 4.5 Konversi Atribut Awal Penerimaan Siswa Baru

No.	Nama Siswa	JK	Jarak Rumah (KM)	Pekerjaan Ayah	Pekerjaan Ibu	Nilai Raport	Status
1	ABDULRAHMAT DUNGGIO	1	2	9	3	86,91	Diterima
2	ALFADLI RAHMAN	1	2	7	1	84,83	Diterima
3	SYAHRIL PAKAYA	1	4	1	1	84,23	Tidak Diterima
4	AQRIL PUTRA NENTO	1	2	3	1	87,83	Diterima
5	ARDIL TAUFIK ABDULLAH	1	3	8	1	85,73	Diterima

No.	Nama Siswa	JK	Jarak Rumah (KM)	Pekerjaan Ayah	Pekerjaan Ibu	Nilai Raport	Status
6	FAREL PONINI	1	2	7	1	87,02	Diterima
7	ISMAIL S HARUN	1	2	3	1	82	Diterima
8	MOH.ALIF FARELYANSYAH LAPAMONA	1	2	9	2	86,53	Diterima
9	MOH.REZKI MARDJUN	1	1	1	1	85,46	Diterima
10	CITRA UMAR	2	5	1	1	81,47	Tidak Diterima
11	MOH.WAHYU KURNIA S.UMAR	1	2	9	1	86,83	Diterima
12	MOH.ADIEL ULYA K TAHIR	1	1	7	3	85,73	Diterima
13	MUHAMAD FADIL HILIMI	1	1	1	1	87,33	Diterima
14	RAMADANI ADAM	1	2	1	1	87,54	Diterima
15	SAHRUL SOLEMAN	1	1	9	1	86,84	Diterima
16	ALFAREL	1	7	5	1	80,56	Tidak Diterima
17	AISYIYAH LIA KOPITOY	2	2	9	1	87,92	Diterima
18	FADHILAH NUR ASYFA S.B	2	1	1	1	86,82	Diterima
.....							
472	ADNAN POHA	1	1	12	6	86.33	Diterima

4.2.2 Pembagian Data Training dan Data Testing

Sebelum dilakukan pemodelan dengan menggunakan metode *K-Nearest Neighbor* terlebih dahulu dilakukan pembagian data menjadi dua yaitu data training dan data testing masing-masing 377:95. Data training digunakan untuk melatih model dan data testing untuk mengukur seberapa baik model dapat menggeneralisasi pada data baru.

Tabel 4.6 Pembagian Data Training dan Data Testing

No	Uraian	Jumlah
1	Data Training	377
2	Data Testing	95
Total		472

4.2.3 Pemodelan dengan *K-Nearest Neighbor*

Tabel 4.7 Sampul Perhitungan Jarak Encludian

No.	Nama Siswa	JK	Jarak Rumah (KM)	Pekerjaan Ayah	Pekerjaan Ibu	Nilai Raport	Status
1	ARDIL TAUFIK ABDULLAH	1	3	8	1	85,73	Diterima
2	FAREL PONINI	1	2	7	1	87,02	Diterima
3	ISMAIL S HARUN	1	2	3	1	82	Diterima
4	CITRA UMAR	2	5	1	1	81,47	Tidak Diterima
5	SYAHRIL PAKAYA	1	4	1	1	84,23	Tidak Diterima
6	MOH.WAHYU KURNIA S.UMAR	1	2	9	1	86,83	?

➤ Berikut ini cara menghitung Jarak Antara Data Training dan Data Testing

$$\begin{aligned}
 \text{Jarak 1} = d1 &= \sqrt{(1-1)^2 + (3-2)^2 + (8-9)^2 + (1-1)^2 + (85,73-86,83)^2} \\
 &= \sqrt{0+1+1+0+1,21} \\
 &= \sqrt{3,21} = 1,7916
 \end{aligned}$$

$$\begin{aligned}
 \text{Jarak 2} = d1 &= \sqrt{(1-1)^2 + (5-2)^2 + (7-9)^2 + (1-1)^2 + (87,02-86,83)^2} \\
 &= \sqrt{0+9+4+0+0,0361} \\
 &= \sqrt{13,0361} = 3.6105
 \end{aligned}$$

$$\begin{aligned}
 \text{Jarak 3} = d1 &= \sqrt{(1-1)^2 + (2-2)^2 + (3-9)^2 + (1-1)^2 + (82-86,83)^2} \\
 &= \sqrt{0+0+36+0+23,3289} \\
 &= \sqrt{59,3289} = 7,7025
 \end{aligned}$$

$$\begin{aligned}
 \text{Jarak 4} = d1 &= \sqrt{(2-1)^2 + (5-2)^2 + (1-9)^2 + (1-1)^2 + (81,47-86,83)^2} \\
 &= \sqrt{1+9+64+0+28,7296} \\
 &= \sqrt{102,7296} = 10,1355
 \end{aligned}$$

$$\begin{aligned}
 \text{Jarak 5} = d1 &= \sqrt{(1-1)^2 + (4-2)^2 + (1-9)^2 + (1-1)^2 + (84,23-86,83)^2} \\
 &= \sqrt{0+4+64+0+6,76} \\
 &= \sqrt{74,76} = 8,6463
 \end{aligned}$$

Berdasarkan hasil perhitungan jarak terhadap 5 data training dengan menggunakan nilai $K = 3$ maka didapatkan nilai terkecil yaitu data ke 1 = 1,7916 data ke 2 = 3,6105 dan data ke 3 = 7.7025, dimana ketiga tetangga terdekat tersebut memiliki status = Ya maka dapat disimpulkan bahwa data baru yang di uji dapat diprediksi statusnya adalah Ya.

4.3. Pengujian Model Dan Evaluasi Kinerja Model

Setelah dilakukan pemodelan prediksi siswa baru dengan metode *K-NN* maka perlu dilakukan pengujian model dengan menggunakan akurasi score dan evaluasi kinerja model menggunakan *Confusion Matrix* yang terdiri dari akurasi, recall, precision, f1-score dan support.

4.3.1 Implementasi Hasil Pemodelan

Pada sub bab ini akan menjelaskan tentang tahapan implementasi dari uji coba Pemodelan untuk Prediksi Penerimaan siswa baru

4.3.2 Implementasi Data Processing

Sebelum dilakukan pemodelan ke dalam metode *K-Nearest Neighbor* terlebih dahulu dilakukan beberapa pengolahan data sebagai berikut :

1. Pseudocode Import Library Python

```

1  import pandas as pd
2  import numpy as np
3  import seaborn as sns
4  import matplotlib.pyplot as plt
5  from google.colab import drive
6  from sklearn.model_selection import train_test_split
7  from sklearn.neighbors import KNeighborsClassifier
8  from sklearn.metrics import accuracy_score
9  from sklearn.metrics import confusion_matrix,
10  classification_report
11
12  %matplotlib inline
13  from joblib import dump, load

```

Pesudcode 4.1 Import Library Python

Penjelasan mengenai library di atas ditunjukkan pada Tabel 4.1

Nama Library	Deskripsi
--------------	-----------

Pandas	mengimpor pustaka Pandas dengan alias pd. Pandas adalah pustaka yang digunakan untuk manipulasi dan analisis data
Numpy	mengimpor pustaka NumPy dengan alias np. NumPy adalah pustaka yang digunakan untuk komputasi numerik dalam Python.
Seabron	mengimpor pustaka Seaborn dengan alias sns. Seaborn adalah pustaka yang digunakan untuk visualisasi data statistik.
matplotlib	mengimpor modul pyplot dari pustaka Matplotlib dengan alias plt. Matplotlib adalah pustaka untuk membuat visualisasi grafik dalam Python.
Google colab	mengimpor modul drive dari pustaka google.colab. Google Colab adalah platform cloud yang memungkinkan pengguna untuk menulis dan mengeksekusi kode Python di browser web.
sklearn.model_selection	mengimpor fungsi train_test_split dari pustaka scikit-learn. Fungsi ini digunakan untuk membagi dataset menjadi subset pelatihan dan pengujian.
sklearn.neighbors	mengimpor kelas KNeighborsClassifier dari pustaka scikit-learn. KNeighborsClassifier adalah model pembelajaran mesin untuk klasifikasi berdasarkan k-Nearest Neighbors.
sklearn.metrics import accuracy_score	mengimpor fungsi accuracy_score dari pustaka scikit-learn. Fungsi ini digunakan untuk menghitung akurasi prediksi dari model pembelajaran mesin.
sklearn.metrics import confusion_matrix, classification_report	mengimpor fungsi confusion_matrix dan classification_report dari pustaka scikit-learn. Fungsi

	ini digunakan untuk mengevaluasi kinerja model pembelajaran mesin.
--	--

2. Pseudocode menghubungkan Google Colab dengan Google Drive

1	<code>from google.colab import drive</code>
2	<code>drive.mount('/content/drive')</code>

Kode tersebut adalah untuk mengakses Google Drive dari Google Colab, yang merupakan platform cloud yang memungkinkan Anda menulis dan mengeksekusi kode Python di browser web.

3. pseudocode Load Dataset

1	<code>#Baca Dataset</code>
2	<code>data =</code>
3	<code>pd.read_excel('/content/drive/MyDrive/skripsi_nurmeilan/dat</code>
	<code>asiswa.xlsx')</code>
4	<code>data</code>

Kode tersebut bertujuan untuk membaca dataset yang tersimpan dalam format Excel dari Google Drive menggunakan Pandas, dan menampilkannya di dalam notebook Google Colab.

4. Pseudocode konversi dataset

1	<code>#Konversi dataset Excel ke CSV</code>
2	<code>data.to_csv("/content/drive/MyDrive/skripsi_nurmeilan/datas</code>
	<code>iswa.csv", index=False)</code>
3	<code>data.info()</code>

Coding tersebut bertujuan untuk mengkonversi dataset yang telah dibaca sebelumnya dari format Excel menjadi format CSV, serta memberikan informasi tentang dataset tersebut.

5. Pseudocode untuk menyimpan data

```

1 data=
2 pd.read_csv('/content/drive/MyDrive/skripsi_nurmeilan/datasi
3 swa.csv')
  data.head()

```

Coding tersebut bertujuan untuk membaca dataset yang telah dikonversi sebelumnya dari format Excel menjadi format CSV, dan menampilkan beberapa baris pertama dari dataset tersebut.

6. Pseudocode data jenis kelamin

```

1 data['Jenis Kelamin'].replace(['L','P'],[1,2], inplace=True)
2 data.head()
3

```

Coding tersebut bertujuan untuk mengganti nilai dalam kolom 'Jenis Kelamin' dengan nilai baru yang telah ditentukan.

7. Pseudocode data pekerjaan ayah

```

1 data['Pekerjaan
2 Ayah'].replace(['Buruh','Petani','Nelayan','Tukang','Sopir','
3 Honorer','Karyawan Swasta',
4 'Pedagang',
5 'Wiraswasta','Pensiunan','TNI/Polri','PNS','Pekerjaan
6 Lainnya'],
7 [1,2,3,4,5,6,7,8,9,10,11,12,13
8 ], inplace=True)
  data.head()

```

Coding tersebut bertujuan untuk mengganti nilai dalam kolom 'Pekerjaan Ayah' dengan nilai numerik yang baru sesuai dengan kategorinya.

8. Pseudocode data pekerjaan ibu

1	<code>data['Pekerjaan Ibu'].replace(['IRT', 'Karyawan Honorar', 'Karyawan Swasta', 'Guru', 'Wiraswasta', 'PNS'], [1,2,3,4,5,6],</code>
2	<code>inplace=True)</code>
3	<code>data.head()</code>

Kode tersebut bertujuan untuk mengganti nilai dalam kolom 'Pekerjaan Ibu' dengan nilai numerik yang baru sesuai dengan kategorinya.

9. Pseudocode mengecek jumlah atribut yang kosong

1	<code>#Mengecek Jumlah Atribut yg kosong</code>
2	<code>data.isnull().sum()</code>

Kode tersebut digunakan untuk menghitung jumlah nilai-nilai kosong (NULL atau NaN) dalam setiap kolom dari suatu DataFrame yang disebut `data`

10. Pseudocode menyimpan data

1	<code>#Menyimpan dataset yang sudah fix untuk di training</code>
2	<code>data.to_csv("/content/drive/MyDrive/skripsi_nurmeilan/dataset siswa.csv", index=False)</code>
3	<code>data.to_excel("/content/drive/MyDrive/skripsi_nurmeilan/datas etsiswa.xlsx", index=False)</code>

Kode tersebut digunakan untuk menyimpan DataFrame yang disebut `data` ke dalam format file CSV dan Excel.

11. Pseudocode untuk menampilkan data

1	<code>data=</code>
2	<code>pd.read_csv('/content/drive/MyDrive/skripsi_nurmeilan/datase tsiswa.csv')</code>
3	<code>data.head()</code>

--	--

Kode tersebut bertujuan untuk membaca sebuah file CSV yang terletak di path `‘/content/drive/MyDrive/skripsi_nurmeilan/datasetsiswa.csv’` dan menampilkannya dalam bentuk Data Frame menggunakan library panda di Python.

12. Pseudocode dataset

1	<code>X = data.iloc[:,2:7] #independent columns</code>
2	<code>y = data.iloc[:, -1] #target column status penerimaan</code>

Coding tersebut bertujuan untuk membagi dataset menjadi dua bagian: satu untuk fitur atau variabel independen (X) dan yang lainnya untuk target atau variabel dependen (y).

13. Pseudocode input data

1	<code>X = data[['Jenis Kelamin', 'Jarak Rumah', 'Pekerjaan Ayah', 'Pekerjaan Ibu', 'Nilai Raport']].values</code>
2	<code>#variabel input</code>
3	<code>y = data[['Status']].values #varibael target/output</code>

Kode tersebut bertujuan untuk menyiapkan variabel input (X) dan variabel target/output (y) dari sebuah DataFrame yang disebut `data`

14. Pseudocode data training dan data testing

1	<code># Train test split untuk membagi data training dan testing</code>
---	---

```

2  from sklearn.model_selection import train_test_split
3
4  X_train, X_test, y_train, y_test = train_test_split(X, y,
5  test_size=0.2, random_state=10)
6
7  print ('Train set:', X_train.shape,  y_train.shape)
8  print ('Test set:', X_test.shape,  y_test.shape)

```

Kode tersebut digunakan untuk membagi dataset menjadi dua bagian: satu untuk data training dan yang lainnya untuk data testing, menggunakan metode train-test split dari library scikit-learn.

15. Pseudocode Model_knn

```

1  k = 3
2  # Train Model
3  model_knn = KNeighborsClassifier(n_neighbors =
4  k).fit(X_train, y_train)
5  model_knn

```

Kode tersebut bertujuan untuk melatih model *K-Nearest Neighbors (KNN)* menggunakan dataset yang telah dibagi sebelumnya menjadi data training (X_{train} , y_{train}) dan data testing (X_{test} , y_{test}).

16. Pseudocode simpan model ke file

```

1  # simpan model ke file
2  dump(model_knn,
3  '/content/drive/MyDrive/skripsi_nurmeilan/knn_model.joblib')

```

Baris kode ini bertujuan untuk menyimpan model yang telah dilatih ke dalam file menggunakan joblib, sebuah modul dalam Python yang berguna untuk menyimpan dan memuat objek Python ke dalam file.

17. Pseudocode menguji model dengan data testing

1	# Menguji model dengan data testing
2	y_pred = model_knn.predict(X_test)
3	y_pred[0:5]

Baris kode ini bertujuan untuk menguji model yang telah dilatih menggunakan data testing (**X_test**) dan memperoleh prediksi dari model untuk data testing tersebut.

18. Pseudocode menampilkan data testing

1	# Menampilkan data testing
2	y_test[0:5]

Baris kode tersebut bertujuan untuk menampilkan label/target yang sebenarnya dari data testing. Ini berguna untuk membandingkan dengan prediksi yang dihasilkan oleh model, sehingga kita dapat mengevaluasi seberapa baik model tersebut melakukan prediksi.

19. Pseudocode *confusion matrix*

1	from sklearn.metrics import confusion_matrix, classification_report
2	import seaborn as sns
3	import matplotlib.pyplot as plt
4	# Evaluasi Model
5	y_pred = model_knn.predict(X_test)
6	# Membuat confusion matrix
7	cm = confusion_matrix(y_test, y_pred)
8	# Mencetak confusion matrix

```

9 labels = ['TIDAK', 'YA']
10 plt.figure(figsize=(5, 4))
11 sns.heatmap(cm, annot=True, fmt='d', xticklabels=labels,
12 yticklabels=labels, cmap='Blues')
13 plt.xlabel('Prediksi')
14 plt.ylabel('Actual')
15 plt.title('Confusion Matrix')
16 plt.show()
17 # Menghitung precision, recall, f1-score, dan support
18 report = classification_report(y_test, y_pred,
19 target_names=labels)
20 print(report)

```

Kode di atas digunakan untuk mengevaluasi model klasifikasi, seperti model *K-Nearest Neighbors (KNN)*, dengan menggunakan beberapa metrik evaluasi, yaitu confusion matrix, precision, recall, dan *f1 score*

4.3.3 Implementasi Pembagian Data Training dan Data Testing

```

1 X = data[['Jenis Kelamin', 'Jarak Rumah', 'Pekerjaan Ayah',
2         'Pekerjaan Ibu', 'Nilai Raport']].values
3 #variabel input
4 y = data[['Status']].values #varibael target/output

```

Kode yang di atas bertujuan untuk mengekstrak variabel input (fitur) dan variabel target (label) dari sebuah dataset yang disimpan dalam objek data

```

1 # Train test split untuk membagi data training dan testing
2 from sklearn.model_selection import train_test_split
3 X_train, X_test, y_train, y_test = train_test_split(X, y,
4 test_size=0.2, random_state=10)
5 print ('Train set:', X_train.shape, y_train.shape)
6 print ('Test set:', X_test.shape, y_test.shape)

```

--	--

Kode yang di atas bertujuan untuk membagi dataset menjadi dua bagian: data pelatihan (training) dan data pengujian (testing) menggunakan fungsi `train_test_split` dari pustaka `scikit-learn`.

4.3.4 Implementasi Pemodelan

1	<code>k = 3</code>
2	<code># Train Model</code>
3	<code>model_knn = KNeighborsClassifier(n_neighbors =</code> <code>k).fit(X_train, y_train)</code>
4	<code>model_knn</code>

Kode ini digunakan untuk melatih sebuah model *K-Nearest Neighbors* (KNN) menggunakan pustaka `scikit-learn`

1	<code># simpan model ke file</code>
2	<code>dump(model_knn,</code>
3	<code>'/content/drive/MyDrive/skripsi_nurmeilan/knn_model.joblib'</code>

Kode ini digunakan untuk menyimpan model yang telah dilatih ke dalam sebuah file menggunakan pustaka `joblib`

4.3.5 Implementasi Pengujian Model dan Evaluasi Kinerja Model

1	<code># Mengukur kinerja model machine learning</code>
2	<code>from sklearn.metrics import accuracy_score</code>
3	<code>print('Akurasi Train set: ', accuracy_score(y_train,</code> <code>model_knn.predict(X_train)))</code>
4	<code>print('Akurasi Test set: ', accuracy_score(y_test, y_pred))</code>

Kode ini digunakan untuk mengukur kinerja model machine learning dengan menghitung akurasi pada set data pelatihan (training set) dan set data pengujian (test set).

```
1  from sklearn.metrics import confusion_matrix,
    classification_report
2  import seaborn as sns
3  import matplotlib.pyplot as plt
4
5  # Evaluasi Model
6  y_pred = model_knn.predict(X_test)
7
8  # Membuat confusion matrix
9  cm = confusion_matrix(y_test, y_pred)
10
11 # Mencetak confusion matrix
12 labels = ['TIDAK', 'YA']
13 plt.figure(figsize=(5, 4))
14 sns.heatmap(cm, annot=True, fmt='d', xticklabels=labels,
15             yticklabels=labels, cmap='Blues')
16 plt.xlabel('Prediksi')
17 plt.ylabel('Actual')
18 plt.title('Confusion Matrix')
19 plt.show()
20
21 # Menghitung precision, recall, f1-score, dan support
22 report = classification_report(y_test, y_pred,
23                               target_names=labels)
24 print(report)
```

Kode yang diberikan bertujuan untuk mengevaluasi kinerja model klasifikasi dengan menggunakan confusion matrix dan laporan klasifikasi yang mencakup metrik seperti precision, recall, F1-score, dan support.

4.3.6 Implementasi Hasil Prediksi

```

1  #Prediksi dengan data Testing Baru dengan cara upload file
   Excel
2  import pandas as pd
3  from sklearn.neighbors import KNeighborsClassifier
4  from google.colab import files
5  import io
6  import ipywidgets as widgets

7  def load_file(change):
8      uploaded = list(load_button.value.values())[0]
9      if uploaded:
10         # Read uploaded file
11         df = pd.read_excel(io.BytesIO(uploaded['content']))
12
13         df['Jenis Kelamin'].replace(['L', 'P'], [1, 2],
14         inplace=True)

15         df['Pekerjaan
   Ayah'].replace(['Buruh', 'Petani', 'Nelayan', 'Tukang', 'Sopir',
   'Honorar', 'Karyawan Swasta',
16                    'Pedagang',
   'Wiraswasta', 'Pensiunan', 'TNI/Polri', 'PNS', 'Pekerjaan
   Lainnya'],
17                    [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 1
18                    3], inplace=True)

19         df['Pekerjaan Ibu'].replace(['IRT', 'Karyawan
   Honorar', 'Karyawan Swasta', 'Guru', 'Wiraswasta', 'PNS'],
20                    [1, 2, 3, 4, 5, 6], inplace=True)

21         X_test = df[['Jenis Kelamin', 'Jarak
   Rumah', 'Pekerjaan Ayah',
22                    'Pekerjaan Ibu', 'Nilai Raport']].values

23         # Perform prediction
24         predictions = loaded_model_knn.predict(X_test)

```

```
19 # Add prediction column to DataFrame
    df['Prediction'] = predictions

20 # Save DataFrame with predictions to a new Excel
    file
21     output_filename = '/content/drive/MyDrive/Enric
2024/testing_data_with_predictions.xlsx'
    df.to_excel(output_filename, index=False)

22     print("Prediction results have been saved to:",
23 output_filename)

24 # Initialize KNN model with k=3
25 k = 3

26 # Create file upload widget
load_button = widgets.FileUpload(accept='.xlsx',
multiple=False)
```

Kode ini memungkinkan pengguna untuk mengunggah file Excel berisi data baru untuk diprediksi menggunakan *model K-Nearest Neighbors (KNN)* yang telah dilatih sebelumnya.

BAB V

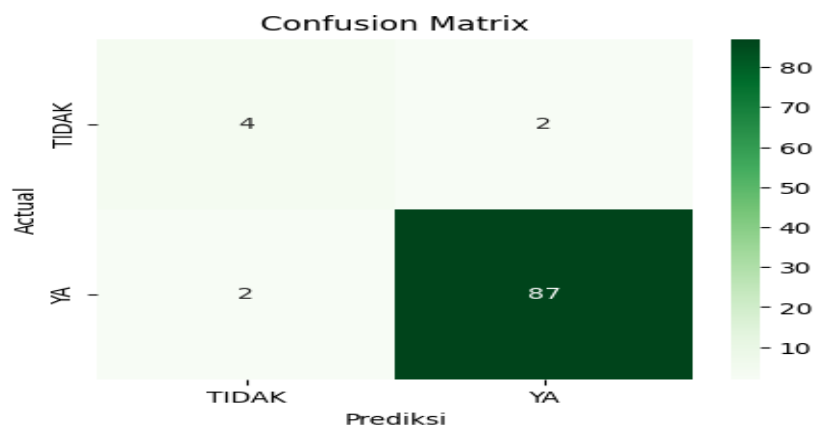
PEMBAHASAN PENELITIAN

5.1 Pembahasan Kinerja Model

Evaluasi kinerja model *K-Nearest Neighbor* dalam prediksi penerimaan siswa baru pada penelitian ini dilakukan uji coba untuk nilai $k=3$, $k=5$ dan $k=7$ dalam mencari nilai kinerja model yang paling optimal. Kinerja model ini akan diukur dari aspek *Precision*, *Recall*, *f1-score* dan *accuracy* dengan menggunakan *confusion matrix*

1. Kinerja Model *K-Nearest Neighbor* dengan nilai $k=3$

Pengukuran kinerja model *K-NN* dalam prediksi penerimaan siswa baru menggunakan data testing sebanyak 95 data untuk nilai $k=3$ Adapun hasil pengukuran kinerjanya ditunjukkan pada Tabel 5.1 berikut :



Gambar 5.1 *Confusion Matrix* model *K-NN* untuk $k=3$

Tabel 5.1 *Classification Report* Model *K-NN* untuk $k=3$

	<i>Precision</i>	<i>Recall</i>	<i>F1-score</i>
Ya	0.98	0.98	0.98
Tidak	0.67	0.67	0.67
Akurasi	0.96		

- Menghitung rumus *Precision*

$$\frac{TP}{TP + FP} = \frac{87}{87 + 2} = \frac{87}{89} = 0,98$$

- Menghitung rumus *Recall*

$$\frac{TP}{TP + FN} = \frac{87}{87 + 2} = \frac{87}{89} = 0,98$$

- Menghitung rumus *f1-score*

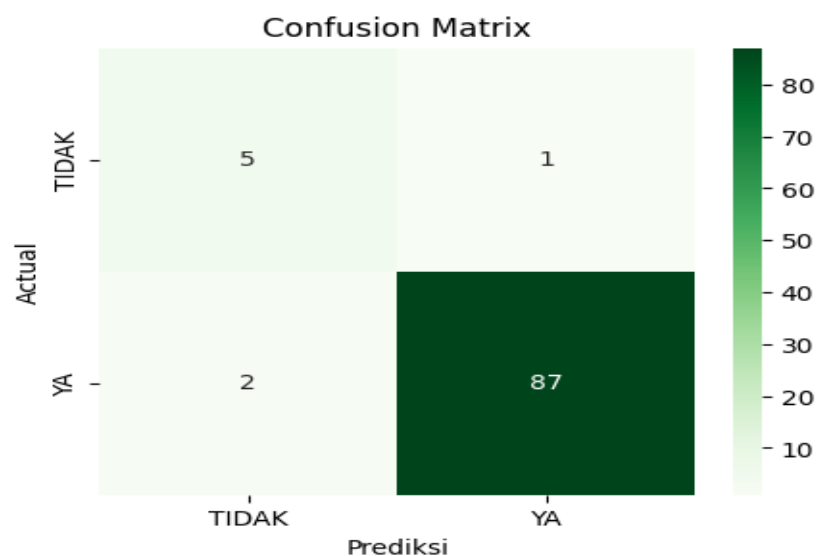
$$2 \times \frac{Precision \times Recall}{Precision + Recall} = 2 \times \frac{0,98 \times 0,98}{0,98 + 0,98} = 2 \times \frac{0,9604}{1,96} = 2 \times 0,49 = 0,98$$

- Menghitung Rumus Akurasi

$$\frac{TP + TN}{TP + TN + FP + FN} = \frac{87 + 2}{87 + 2 + 2 + 2} = \frac{89}{93} = 0,96$$

2. Kinerja Model *K-Nearest Neighbor* dengan nilai $k=5$

Pengukuran kinerja model *K-NN* dalam prediksi penerimaan siswa baru menggunakan data testing sebanyak 95 data untuk nilai $k=5$ Adapun hasil pengukuran kinerjanya ditunjukkan pada Tabel 5.2 berikut :



Gambar 5.2 *Confusion Matrix* model *KNN* untuk $k=5$

Tabel 5.2 *Classification Report Model KNN untuk k=5*

	<i>Precision</i>	<i>Recall</i>	<i>F1-score</i>
Ya	0.99	0.98	0.98
Tidak	0.71	0.83	0.77
Akurasi	0.97		

- Menghitung rumus *Precision*

$$\frac{TP}{TP + FP} = \frac{87}{87 + 1} = \frac{87}{88} = 0,99$$

- Menghitung rumus *Recall*

$$\frac{TP}{TP + FN} = \frac{87}{87 + 2} = \frac{87}{89} = 0,98$$

- Menghitung rumus *f1-score*

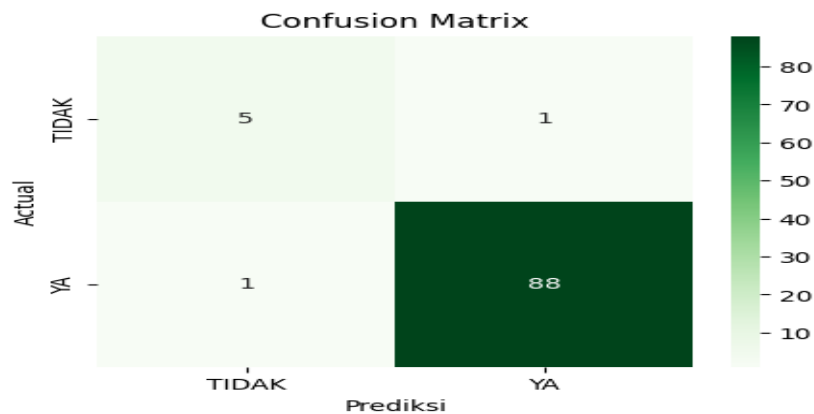
$$2 \times \frac{Precision \times Recall}{Precision + Recall} = 2 \times \frac{0,99 \times 0,98}{0,99 + 0,98} = 2 \times \frac{0,9702}{1,97} = 2 \times 0,49 = 0,98$$

- Menghitung Rumus Akurasi

$$\frac{TP + TN}{TP + TN + FP + FN} = \frac{87 + 5}{87 + 5 + 1 + 2} = \frac{92}{95} = 0,97$$

3. Kinerja Model *K-Nearest Neighbor* dengan nilai k=7

Pengukuran kinerja model *K-NN* dalam prediksi penerimaan siswa baru menggunakan data testing sebanyak 95 data untuk nilai k=7. Adapun hasil pengukuran kinerjanya ditunjukkan pada Tabel 5.3 berikut :



Gambar 5.3 *Confusion Matrix* model KNN untuk k=7

Tabel 5.3 *Classification Report* Model KNN untuk k=7

	<i>Precision</i>	<i>Recall</i>	<i>F1-score</i>
Ya	0.99	0.98	0.98
Tidak	0.83	0.83	0.83
Akurasi	0.98		

- Menghitung rumus *Precision*

$$\frac{TP}{TP + FP} = \frac{88}{88 + 1} = \frac{88}{89} = 0,99$$

- Menghitung rumus *Recall*

$$\frac{TP}{TP + FN} = \frac{88}{88 + 1} = \frac{88}{89} = 0,98$$

- Menghitung rumus *f1-score*

$$2 \times \frac{Precision \times Recall}{Precision + Recall} = 2 \times \frac{0,99 \times 0,98}{0,99 + 0,98} = 2 \times \frac{0,9702}{1,97} = 2 \times 0,49 = 0,98$$

- Menghitung Rumus Akurasi

$$\frac{TP + TN}{TP + TN + FP + FN} = \frac{88 + 5}{88 + 5 + 1 + 1} = \frac{93}{95} = 0,98$$

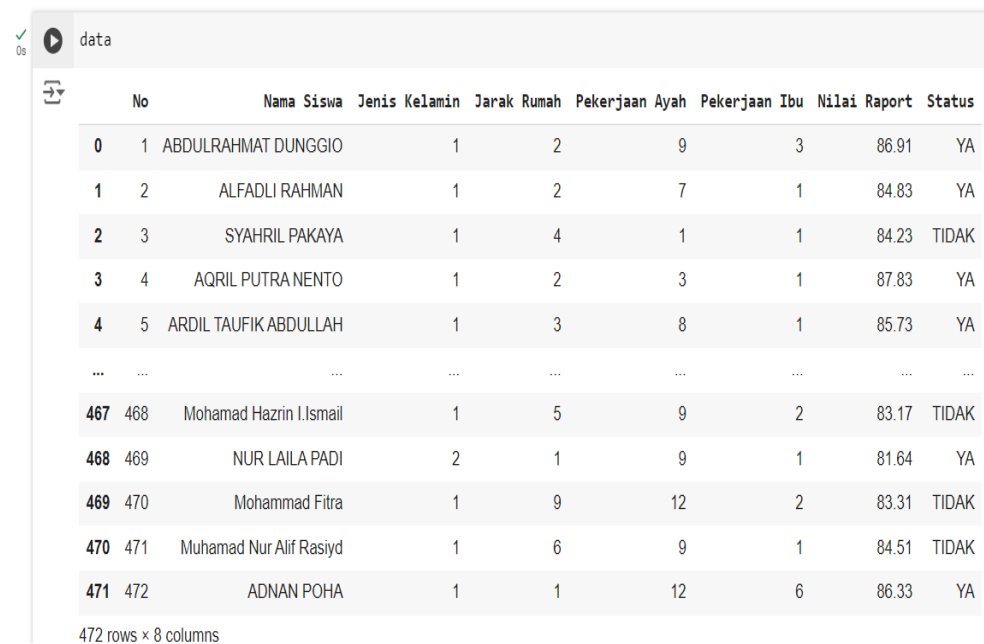
Berdasarkan perhitungan kinerja model *K-NN* yang ditunjukkan pada Tabel 5.1, Tabel 5.2, dan Tabel 5.3 dimana nilai akurasi tertinggi yaitu nilai $k=7$ dengan nilai akurasi 98% sehingga pemodelan *K-NN* dalam prediksi penerimaan siswa baru digunakan nilai $k=7$ sebagai nilai k yang terbaik.

5.2 Pembahasan Hasil Pemodelan

Penerapan metode *K-Nearest Neighbor* dalam prediksi penerimaan siswa baru dengan proses pemodelan dengan tahapan sebagai berikut :

1. Preprocessing

Pada tahapan ini beberapa proses dilakukan diantaranya dataset penerimaan siswa dalam bentuk file excel dikonversi menjadi file csv agar dapat diolah di python lebih cepat dan lebih mudah. Setelah itu dilakukan pengecekan type data atribut yang ada, agar dapat diproses pada metode *K-Nearest Neighbor* semua atribut input memiliki type data yang sama, dimana atribut yang digunakan yaitu Jenis Kelamin, Pekerjaan Ayah, dan Pekerjaan Ibu 3 atribut yang isinya kategorikal sehingga harus diubah menjadi numerik. Adapun hasil penyesuaian ketiga atribut tersebut



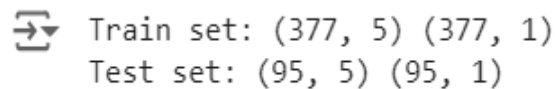
No	Nama Siswa	Jenis Kelamin	Jarak Rumah	Pekerjaan Ayah	Pekerjaan Ibu	Nilai Raport	Status
0 1	ABDULRAHMAT DUNGGIO	1	2	9	3	86.91	YA
1 2	ALFADLI RAHMAN	1	2	7	1	84.83	YA
2 3	SYAHRIL PAKAYA	1	4	1	1	84.23	TIDAK
3 4	AQRIL PUTRA NENTO	1	2	3	1	87.83	YA
4 5	ARDIL TAUFIK ABDULLAH	1	3	8	1	85.73	YA
...	...	...	...	...	...	...	...
467 468	Mohamad Hazrin I. Ismail	1	5	9	2	83.17	TIDAK
468 469	NUR LAILA PADI	2	1	9	1	81.64	YA
469 470	Mohammad Fitra	1	9	12	2	83.31	TIDAK
470 471	Muhamad Nur Alif Rasyid	1	6	9	1	84.51	TIDAK
471 472	ADNAN POHA	1	1	12	6	86.33	YA

472 rows × 8 columns

Gambar 5.4 Hasil Penyesuaian Atribut

2. Pembagian Data Training dan Data Testing

Pembagian Data Training dan Data Testing untuk pemodelan *K-Nearest Neighbor* dalam prediksi penerimaan siswa baru yaitu 80 : 20 sehingga didapatkan masing-masing Data Training dan Data Testing sebagaimana ditunjukkan pada Gambar 5.5 dibawah ini.



Train set: (377, 5) (377, 1)
Test set: (95, 5) (95, 1)

Gambar 5.5 Pembagian Data Training dan Data Testing

3. Pemodelan dengan *K-Nearest Neighbor*

Pemodelan *K-NN* untuk melakukan prediksi Penerimaan siswa baru dilakukan sebanyak tiga kali yaitu nilai $k=3$, nilai $k=5$, dan nilai $k=7$. Berdasarkan hasil pemodelan didapatkan model *KNN* $k=7$ yang terbaik sehingga model ini yang digunakan dalam proses prediksi penerimaan siswa baru

4. Evaluasi Kinerja Model *K-Nearest Neighbor*

Setelah dilakukan pemodelan sebanyak tiga model berikutnya dilakukan evaluasi kerja untuk setiap model dengan mengevaluasi kinerja terdiri dari *Precision*, *Recall*, *f1-score* dan *Accuracy* dimana hasilnya ditunjukkan pada Tabel 5.1, Tabel 5.2, dan Table 5.3

5. Pengujian Model pada Data Testing Baru

Setelah didapatkan model *K-Nearest Neighbor* yang memiliki akurasi tertinggi yaitu $k=7$ dengann akurasi sebesar 98%, maka model tersebut digunakan untuk melakukan prediksi penerimaan siswa baru. Adapun cara melakukan prediksi penerimaan siswa baru yaitu data-data siswa baru dibuat dalam bentuk file excel seperti yang ditunjukkan pada Tabel 5.4

Tabel 5.4 Pengujian data testing baru

No	Nama Siswa	JK	Jarak Rumah (KM)	Pekerjaan Ayah	Pekerjaan Ibu	Nilai Raport
1	ABDULRAHMAT DUNGGIO	L	2	Wiraswasta	Karyawan Swasta	86.91
2	ALFADLI RAHMAN	L	2	Karyawan Swasta	IRT	84.83
3	SYAHRIL PAKAYA	L	4	Buruh	IRT	84.23
4	AQRIL PUTRA NENTO	L	2	Nelayan	IRT	87.83
5	ARDIL TAUFIK ABDULLAH	L	3	Pedagang	IRT	85.73
6	FAREL PONINI	L	5	Karyawan Swasta	IRT	87.02
7	ISMAIL S HARUN	L	2	Nelayan	IRT	82
8	MOH.ALIF FARELYANSYAH LAPAMONA	L	2	Wiraswasta	Karyawan Honorer	86.53
9	MOH.REZKI MARDJUN	L	1	Buruh	IRT	85.46
10	CITRA UMAR	P	5	Buruh	IRT	81.47
11	MOH.WAHYU KURNIA S.UMAR	L	2	Wiraswasta	IRT	86.83
12	MOH.ADIEL ULYA K TAHIR	L	1	Karyawan Swasta	Karyawan Swasta	85.73
13	MUHAMAD FADIL HILIMI	L	1	Buruh	IRT	87.33
14	RAMADANI ADAM	L	2	Buruh	IRT	87.54
15	SAHRUL SOLEMAN	L	1	Wiraswasta	IRT	86.84
16	ALFAREL	L	7	Sopir	IRT	80.56
17	AISYIYAH LIA KOPITOY	P	2	Wiraswasta	IRT	87.92
18	FADHILAH NUR ASYFA S.B	P	1	Buruh	IRT	86.82
19	FAUZIA PUTRI DAENG MAPILE	P	1	Wiraswasta	IRT	87.17
20	FAZRIN SALSABILAH PAKAJA	P	3	PNS	Guru	87.64
21	FIDYA AMALIA DJUNUBI	P	1	PNS	Guru	84.71
22	FIDYA ANGGRAINI DJAFAR	P	2	Wiraswasta	IRT	82.88
23	MARCHIKA S.SUAIB	P	1	PNS	IRT	84.77
24	MISNAWATI ABDUL	P	2	Wiraswasta	IRT	87.71
25	NUR MEYLAN AINUN ALI	P	3	Pedagang	IRT	83.73

Setelah Data siswa baru sudah dalam bentuk file Excel, selanjutnya dijalankan model *K-Nearest Neighbor* prediksi penerimaan siswa baru pada tools Google Colabratory. Hasil prediksi data testing baru dengan menggunakan model *KNN* yang terbaik hasilnya disimpan ke dalam bentuk file excel seperti yang ditunjukkan pada Tabel 5.5

Tabel 5.5 Hasil Pengujian DataTesting Baru

No	Nama Siswa	Jenis Kelamin	Jarak Rumah (KM)	Pekerjaan Ayah	Pekerjaan Ibu	Nilai Raport	Prediction
1	ABDULRAHMAT DUNGGIO	1	2	9	3	86,91	DITERIMA
2	ALFADLI RAHMAN	1	2	7	1	84,83	DITERIMA
3	SYAHRIL PAKAYA	1	4	1	1	84,23	TIDAK DITERIMA
4	AQRIL PUTRA NENTO	1	2	3	1	87,83	DITERIMA
5	ARDIL TAUFIK ABDULLAH	1	3	8	1	85,73	DITERIMA
6	FAREL PONINI	1	5	7	1	87,02	TIDAK DITERIMA
7	ISMAIL S HARUN	1	2	3	1	82	DITERIMA
8	MOH.ALIF FARELYANSYAH LAPAMONA	1	2	9	2	86,53	DITERIMA
9	MOH.REZKI MARDJUN	1	1	1	1	85,46	DITERIMA
10	CITRA UMAR	2	5	1	1	81,47	TIDAK DITERIMA
11	MOH.WAHYU KURNIA S.UMAR	1	2	9	1	86,83	DITERIMA
12	MOH.ADIEL ULYA K TAHIR	1	1	7	3	85,73	DITERIMA
13	MUHAMAD FADIL HILIMI	1	1	1	1	87,33	DITERIMA
14	RAMADANI ADAM	1	2	1	1	87,54	DITERIMA
15	SAHRUL SOLEMAN	1	1	9	1	86,84	DITERIMA
16	ALFAREL	1	7	5	1	80,56	TIDAK DITERIMA
17	AISYIYAH LIA KOPITOY	2	2	9	1	87,92	DITERIMA
18	FADHILAH NUR ASYFA S.B	2	1	1	1	86,82	DITERIMA
19	FAUZIA PUTRI DAENG MAPILE	2	1	9	1	87,17	DITERIMA
20	FAZRIN SALSABILAH PAKAJA	2	3	12	4	87,64	DITERIMA
21	FIDYA AMALIA DJUNUBI	2	1	12	4	84,71	DITERIMA

No	Nama Siswa	Jenis Kelamin	Jarak Rumah (KM)	Pekerjaan Ayah	Pekerjaan Ibu	Nilai Raport	Prediction
22	FIDYA ANGGRAINI DJAFAR	2	2	9	1	82,88	DITERIMA
23	MARCHIKA S.SUAIB	2	1	12	1	84,77	DITERIMA
24	MISNAWATI ABDUL	2	2	9	1	87,71	DITERIMA
25	NUR MEYLAN AINUN ALI	2	3	8	1	83,73	DITERIMA

BAB VI

KESIMPULAN DAN SARAN

6.1 Kesimpulan

Berdasarkan hasil implementasi metode *K-Nearest Neighbor* dalam prediksi penerimaan siswa baru diperoleh kesimpulan sebagai berikut :

1. Penerapan metode *K-Nearest Neighbor* dalam prediksi penerimaan siswa baru dapat diterapkan secara efektif
2. Hasil akurasi penerapan metode *K-Nearest Neighbor* dalam prediksi penerimaan siswa baru diperoleh akurasi sebesar 98% dengan nilai k (tentangga terdekat) terbaik yaitu $k = 7$

6.2 Saran

Berdasarkan kesimpulan diatas, saran bagi peneliti berikutnya yaitu :

1. Dapat menambahkan atribut yang lain berkaitan dengan penerimaan siswa baru agar hasil prediksinya bisa disesuaikan dengan kondisi tempat objek penelitian
2. Dapat menggunakan metode prediksi yang lain agar dapat dibandingkan hasilnya mana yang lebih baik

DAFTAR PUSTAKA

- [1] M. Faris, Y. A. Pranoto, and H. Z. Zahro, “Penentuan Penerima Bantuan Sosial Bagi Siswa Yang Terkena Dampak Covid-19 Menggunakan Metode K-Nearest Neighbor,” *J. Mhs. Tek. Inform.*, vol. 5, no. 1, pp. 276–283, 2021.
- [2] Nizarman, “Manajemen Penerimaan Siswa Baru,” *Manajer Pendidik.*, vol. 9, no. 2, pp. 224–234, 2015, [Online]. Available: <https://ejournal.unib.ac.id/index.php/manajerpendidikan/article/download/1116/925>
- [3] D. Kurniawan and A. Saputra, “Penerapan K-Nearest Neighbour dalam Penerimaan Peserta Didik dengan Sistem Zonasi,” *J. Sist. Inf. Bisnis*, vol. 9, no. 2, p. 212, 2019, doi: 10.21456/vol9iss2pp212-219.
- [4] W. Apriansyah, S. Bahri, and F. Saputra, “Sistem Informasi Penerimaan Siswa Baru Pada Smp Negeri 3 Air Gegas Berbasis Web,” *Inser. Inf. Syst. Emerg. Technol. J.*, vol. 3, no. 1, pp. 42–52, 2018, [Online]. Available: <https://repository.atmaluhur.ac.id/handle/123456789/1851>
- [5] N. Syamsiyah and I. Tofany, “Rancang Bangun Sistem Informasi Prediksi Pinjaman Pada Koperasi Panca Bhakti Bekasi Menggunakan Algoritma C4.5,” *J. Sains Teknol.*, vol. 9, no. 1, pp. 28–43, 2019.
- [6] A. L. Maukar, F. Marisa, and A. A. Widodo, “Analisis Data Penerimaan Mahasiswa Baru Berbasis K-Means,” *JIKO (Jurnal Inform. dan Komputer)*, vol. 6, no. 2, p. 142, 2022, doi: 10.26798/jiko.v6i2.558.
- [7] A. Nihcolson, D. S. Naga, and V. C. Mawardi, “Penerapan Metode K-Nearest Neighbor Dalam Memprediksi Waktu Kelulusan Mahasiswa Sarjana Yang Bermain Game,” *J. Ilmu Komput. dan Sist. Inf.*, vol. 9, no. 2, p. 55, 2021, doi: 10.24912/jiksi.v9i2.13107.

LAMPIRAN 1 PROGRAM

```
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
```

```
%matplotlib inline
from joblib import dump, load

from google.colab import drive
drive.mount('/content/drive')
```

```
#Baca Dataset
data =
pd.read_excel('/content/drive/MyDrive/skripsi_nurmeilan/datasi
swa.xlsx')
data
```

```
#Konversi dataset Excel ke CSV
data.to_csv("/content/drive/MyDrive/skripsi_nurmeilan/datasw
a.csv", index=False)
data.info()
```

```
data=
pd.read_csv('/content/drive/MyDrive/skripsi_nurmeilan/datasw
a.csv')
data.head()
```

```
data['Status'].value_counts()
```

```
data['Jenis Kelamin'].replace(['L', 'P'], [1, 2], inplace=True)
data.head()
```

```
data['Pekerjaan Ayah'].value_counts()
```

```
data['Pekerjaan
Ayah'].replace(['Buruh', 'Petani', 'Nelayan', 'Tukang', 'Sopir', 'H
onorer', 'Karyawan Swasta',
                'Pedagang',
                'Wiraswasta', 'Pensiunan', 'TNI/Polri', 'PNS', 'Pekerjaan
Lainnya'],
               [1,2,3,4,5,6,7,8,9,10,11,12,13]
, inplace=True)
data.head()
```

```
data['Pekerjaan Ibu'].value_counts()
```

```
data['Pekerjaan Ibu'].replace(['IRT', 'Karyawan
Honoror', 'Karyawan Swasta', 'Guru', 'Wiraswasta', 'PNS'],
                             [1,2,3,4,5,6], inplace=True)
data.head()
```

```
data.describe()
```

```
#Mengecek Jumlah Atribut yg kosong
data.isnull().sum()
```

```
data
```

```
#Menyimpan dataset yang sudah fix untuk di training
data.to_csv("/content/drive/MyDrive/skripsi_nurmeilan/datasets
iswa.csv", index=False)
data.to_excel("/content/drive/MyDrive/skripsi_nurmeilan/datase
tsiswa.xlsx", index=False)
```

```
data=
pd.read_csv('/content/drive/MyDrive/skripsi_nurmeilan/datasets
iswa.csv')
data.head()
```

```
X = data.iloc[:,2:7] #independent columns
y = data.iloc[:, -1] #target column status penerimaan
```

```

X

Y

X = data[['Jenis Kelamin','Jarak Rumah','Pekerjaan Ayah',
          'Pekerjaan Ibu','Nilai Raport']].values #variabel
input
y = data[['Status']].values #varibael target/output

```

```

y

```

```

# Train test split untuk membagi data training dan testing

from sklearn.model_selection import train_test_split

X_train, X_test, y_train, y_test = train_test_split(X, y,
test_size=0.2, random_state=10)

print ('Train set:', X_train.shape,  y_train.shape)
print ('Test set:', X_test.shape,  y_test.shape)

```

```

from sklearn.neighbors import KNeighborsClassifier

```

```

k = 3

# Train Model
model_knn = KNeighborsClassifier(n_neighbors = k).fit(X_train,
y_train)
model_knn

```

```

# simpan model ke file
dump(model_knn,
'/content/drive/MyDrive/skripsi_nurmeilan/knn_model.joblib')

```

```

# Menguji model dengan data testing

y_pred = model_knn.predict(X_test)
y_pred[0:5]

```

```
# Menampilkan data testing
```

```
y_test[0:5]
```

```
# Mengukur kinerja model machine learning
```

```
from sklearn.metrics import accuracy_score
```

```
print('Akurasi Train set: ', accuracy_score(y_train,  
model_knn.predict(X_train)))
```

```
print('Akurasi Test set: ', accuracy_score(y_test, y_pred))
```

```
# confusion matrix untuk k=3
```

```
from sklearn.metrics import confusion_matrix,
```

```
classification_report
```

```
import seaborn as sns
```

```
import matplotlib.pyplot as plt
```

```
# Evaluasi Model
```

```
y_pred = model_knn.predict(X_test)
```

```
# Membuat confusion matrix
```

```
cm = confusion_matrix(y_test, y_pred)
```

```
# Mencetak confusion matrix
```

```
labels = ['TIDAK', 'YA']
```

```
plt.figure(figsize=(5, 4))
```

```
sns.heatmap(cm, annot=True, fmt='d', xticklabels=labels,  
yticklabels=labels, cmap='Greens')
```

```
plt.xlabel('Prediksi')
```

```
plt.ylabel('Actual')
```

```
plt.title('Confusion Matrix')
```

```
plt.show()
```

```
# Menghitung precision, recall, f1-score, dan support
```

```
report = classification_report(y_test, y_pred,  
target_names=labels)
```

```
print(report)
```

```

k = 5

# Train Model
model_knn5 = KNeighborsClassifier(n_neighbors =
k).fit(X_train, y_train)
model_knn5

# confusion matrix untuk k=5
from sklearn.metrics import confusion_matrix,
classification_report
import seaborn as sns
import matplotlib.pyplot as plt

# Evaluasi Model
y_pred = model_knn5.predict(X_test)

# Membuat confusion matrix
cm = confusion_matrix(y_test, y_pred)

# Mencetak confusion matrix
labels = ['TIDAK', 'YA']
plt.figure(figsize=(5, 4))
sns.heatmap(cm, annot=True, fmt='d', xticklabels=labels,
yticklabels=labels, cmap='Greens')
plt.xlabel('Prediksi')
plt.ylabel('Actual')
plt.title('Confusion Matrix')
plt.show()

# Menghitung precision, recall, f1-score, dan support
report = classification_report(y_test, y_pred,
target_names=labels)
print(report)

k = 7

# Train Model
model_knn7 = KNeighborsClassifier(n_neighbors =
k).fit(X_train, y_train)
model_knn7

```

```

# confusion matrix untuk k=7
from sklearn.metrics import confusion_matrix,
classification_report
import seaborn as sns
import matplotlib.pyplot as plt

# Evaluasi Model
y_pred = model_knn7.predict(X_test)

# Membuat confusion matrix
cm = confusion_matrix(y_test, y_pred)

# Mencetak confusion matrix
labels = ['TIDAK', 'YA']
plt.figure(figsize=(5, 4))
sns.heatmap(cm, annot=True, fmt='d', xticklabels=labels,
yticklabels=labels, cmap='Greens')
plt.xlabel('Prediksi')
plt.ylabel('Actual')
plt.title('Confusion Matrix')
plt.show()

# Menghitung precision, recall, f1-score, dan support
report = classification_report(y_test, y_pred,
target_names=labels)
print(report)

# simpan model ke file
dump(model_knn7,
'/content/drive/MyDrive/skripsi_nurmeilan/knn_model.joblib')

```

LAMPIRAN 2 SURAT TELAH MELAKUKAN PENELITIAN

PEMERINTAH KOTA GORONTALO
DINAS PENDIDIKAN
SMP NEGERI 4 GORONTALO

Jln. Beringin No.50 Kel. Buladu Kec. Kota Barat Telp (0435) 824281

SURAT KETERANGAN SELESAI MENELITI

Nomor : 800/Disdik/SMP-4/107/2024

Yang bertanda tangan dibawah ini Kepala SMP Negeri 4 Gorontalo, dengan ini menerangkan kepada :

Nama : NUR MEYLAN PAKAYA

N I M : T3120031

Program Studi : TEKNIK INFORMATIKA

Bahwa berdasarkan surat dari Dinas Penanaman Modal dan Pelayanan Terpadu Satu Pintu Kota Gorontalo Nomor : 4710/PIP/LEMLIT-UNISAN/GTO/IX/2023, Perihal Permohonan Penelitian, maka dengan ini menyatakan bahwa yang bersangkutan telah selesai melakukan penelitian dengan judul : **"IMPLEMENTASI METODE KNN UNTUK PREDIKSI PENERIMAAN SISWA BARU (STUDI KASUS SMP NEGERI 4 KOTA GORONTALO)".**

Demikian surat keterangan ini dibuat untuk dipergunakan sebagaimana mestinya.

Gorontalo, 22 Mei 2024
Plt. Kepala Sekolah

HASANA U. MENU, M.Pd
NIP. 19651025 198601 2 003

LAMPIRAN 3 SURAT KETERANGAN BEBAS PUSTAKA

**KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN TEKNOLOGI**
UNIVERSITAS ICHSAN GORONTALO
FAKULTAS ILMU KOMPUTER
UPT. PERPUSTAKAAN FAKULTAS
SK. MENDIKNAS RI NO. 84/D/0/2001
Jl. Achmad Nadjamuddin No.17 Telp(0435) 829975 Fax. (0435) 829976 Gorontalo

SURAT KETERANGAN BEBAS PUSTAKA
No : 007/Perpustakaan-Fikom/V/2024

Perpustakaan Fakultas Ilmu Komputer (FIKOM) Universitas Ichsan Gorontalo dengan ini menerangkan bahwa :

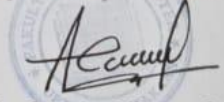
Nama Anggota : Nurmeylan H. Pakaya
No. Induk : T3120031
No. Anggota : M202413

Terhitung mulai hari, tanggal : Sabtu, 25 Mei 2024, dinyatakan telah bebas pinjam buku dan koleksi perpustakaan lainnya.

Demikian keterangan ini di buat untuk di pergunakan sebagaimana mestinya.



Gorontalo, 25 Mei 2024
Mengetahui,
Kepala Perpustakaan



Apriyanto Alhamad, M.Kom
NIDN : 0924048601

LAMPIRAN 4 SURAT REKOMENDASI BEBAS PLAGIASI



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN TEKNOLOGI
UNIVERSITAS ICHSAN GORONTALO
FAKULTAS ILMU KOMPUTER
SURAT KEPUTUSAN MENDIKNAS RI NOMOR 84/D/O/2001
 Jl. Achmad Najamuddin No. 17 Telp. (0435) 829975 Fax (0435) 829976 Gorontalo

SURAT REKOMENDASI BEBAS PLAGIASI
No. 101/FIKOM-UIG/R/V/2024

Yang bertanda tangan di bawah ini :

Nama : Irvan Abraham Salihi, M.Kom
 NIDN : 0928028101
 Jabatan : Dekan Fakultas Ilmu Komputer

Dengan ini menerangkan bahwa :

Nama Mahasiswa : Nurmeylan H. Pakaya
 NIM : T3120031
 Program Studi : Teknik Informatika (S1)
 Fakultas : Fakultas Ilmu Komputer
 Judul Skripsi : Penerapan Metode K-Nearest Neighbor Untuk Prediksi
 Penerimaan Siswa Baru

Sesuai hasil pengecekan tingkat kemiripan skripsi melalui aplikasi **Turnitin** untuk judul skripsi di atas diperoleh hasil *Similarity* sebesar **7%**, berdasarkan Peraturan Rektor No. 32 Tahun 2019 tentang Pendeteksian Plagiat pada Setiap Karya Ilmiah di Lingkungan Universitas Ichsan Gorontalo dan persyaratan pemberian surat rekomendasi verifikasi calon wisudawan dari LLDIKTI Wil. XVI, bahwa batas kemiripan skripsi maksimal 30%, untuk itu skripsi tersebut di atas dinyatakan **BEBAS PLAGIASI** dan layak untuk diujikan.

Demikian surat rekomendasi ini dibuat untuk digunakan sebagaimana mestinya.



Mengetahui
Dekan,
Irvan Abraham Salihi, M.Kom
NIDN. 0928028101

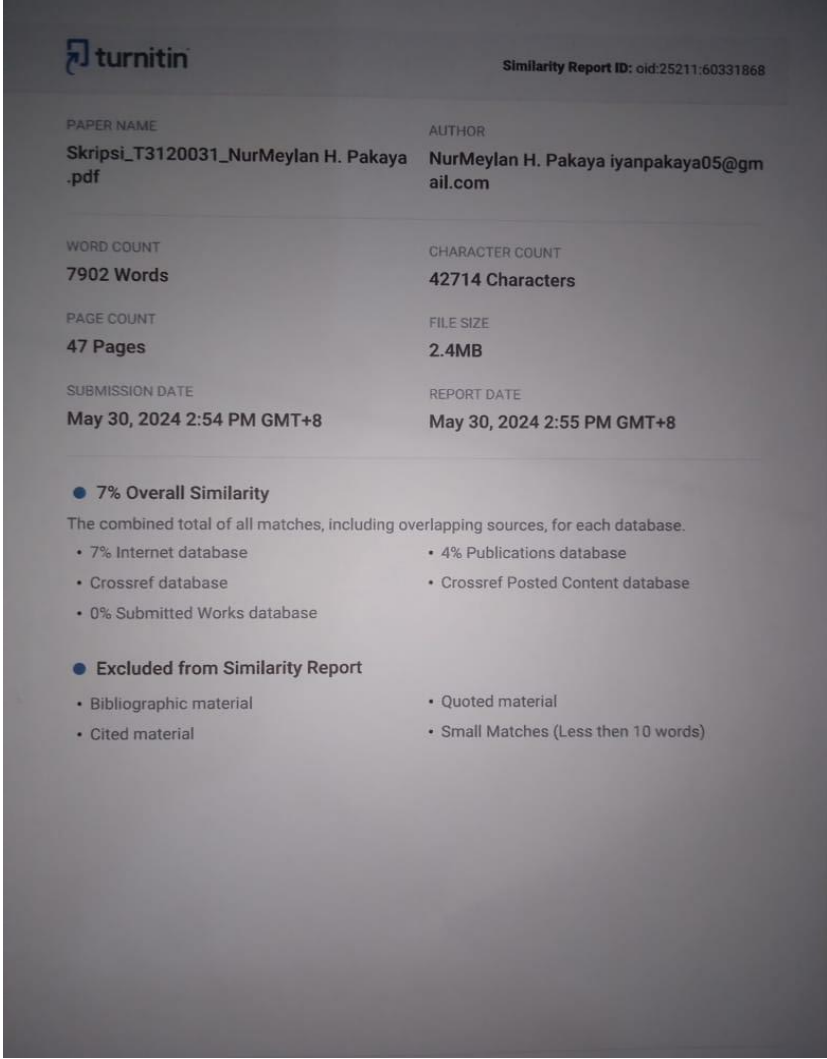
Gorontalo, 30 Mei 2024
Tim Verifikasi,



Zulfrianto Y. Lamasiqi, M.Kom
NIDN. 0914099101

Terlampir :
Hasil Pengecekan Turnitin

LAMPIRAN 5 HASIL TURNITIN



The image is a screenshot of a Turnitin Similarity Report. At the top left is the Turnitin logo. At the top right is the report ID: 'Similarity Report ID: oid:25211:60331868'. The report is divided into two columns. The left column contains: 'PAPER NAME' (Skripsi_T3120031_NurMeylan H. Pakaya .pdf), 'WORD COUNT' (7902 Words), 'PAGE COUNT' (47 Pages), and 'SUBMISSION DATE' (May 30, 2024 2:54 PM GMT+8). The right column contains: 'AUTHOR' (NurMeylan H. Pakaya iyanpakaya05@gmail.com), 'CHARACTER COUNT' (42714 Characters), 'FILE SIZE' (2.4MB), and 'REPORT DATE' (May 30, 2024 2:55 PM GMT+8). Below this, there are two sections: '7% Overall Similarity' and 'Excluded from Similarity Report'. The '7% Overall Similarity' section includes a sub-header 'The combined total of all matches, including overlapping sources, for each database.' and a list of sources: 7% Internet database, 4% Publications database, Crossref database, Crossref Posted Content database, and 0% Submitted Works database. The 'Excluded from Similarity Report' section includes a list of exclusions: Bibliographic material, Cited material, Quoted material, and Small Matches (Less than 10 words).

turnitin Similarity Report ID: oid:25211:60331868

PAPER NAME Skripsi_T3120031_NurMeylan H. Pakaya .pdf	AUTHOR NurMeylan H. Pakaya iyanpakaya05@gmail.com
WORD COUNT 7902 Words	CHARACTER COUNT 42714 Characters
PAGE COUNT 47 Pages	FILE SIZE 2.4MB
SUBMISSION DATE May 30, 2024 2:54 PM GMT+8	REPORT DATE May 30, 2024 2:55 PM GMT+8

● **7% Overall Similarity**

The combined total of all matches, including overlapping sources, for each database.

- 7% Internet database
- 4% Publications database
- Crossref database
- Crossref Posted Content database
- 0% Submitted Works database

● **Excluded from Similarity Report**

- Bibliographic material
- Quoted material
- Cited material
- Small Matches (Less than 10 words)

LAMPIRAN 6 DAFTAR RIWAYAT HIDUP



NurMeylan H. Pakaya, Lahir di momalia pada tanggal 01 mei tahun 2002, anak kedua dari pasangan Bapak Haris pakaya(Alm) dan ibu Suriyati Salilama

Agama : Islam

Jenjang : Strata 1(S1) Sarjana Komputer

Fakultas : Ilmu Komputer Universitas Ichsan
Gorontalo

Alamat : Desa Momalia, Kec.Posigadan Kab. Bolsel

Riwayat Pendidikan

1. Tahun 2014, Menyelesaikan Pendidikan di Sekolah Dasar Negeri 3 Momalia, Kec.Posigadan Kab.Bolaang Mongondow Selatan
2. Tahun 2017, Menyelesaikan Pendidikan di Sekolah Menengah Pertama Negeri 1 Momalia, Kec.Posigadan Kab.Bolaang Mongondow Selatan
3. Tahun 2020, Menyelesaikan Pendidikan di SMK N 1 Posigadan Kab.Bolaang Mongondow Selatan
4. Tahun 2020, Mendaftar dan diterima menjadi Mahasiswa di Fakultas Ilmu Komputer Universitas Ichsan Gorontalo

