

**PENERAPAN ALGORITMA *K-NEAREST NEIGHBOR*
(K-NN) UNTUK PREDIKSI LAMA MASA STUDI
MAHASISWA DI UNIVERSITAS ICHSAN
GORONTALO**

Oleh

FARUQ NATSIR DJAFAR

T3117023

SKRIPSI

Untuk memperoleh salah satu syarat

Guna memperoleh gelar sarjana



PROGRAM SARJANA

TEKNIK INFORMATIKA

UNIVERSITAS ICHSAN GORONTALO

GORONTALO

2022

PENGESAHAN SKRIPSI

**PENERAPAN ALGORITMA *K-NEAREST NEIGHBOR*
(K-NN) UNTUK PREDIKSI LAMA MASA STUDI
MAHASISWA DI UNIVERSITAS ICHSAN
GORONTALO**

Oleh

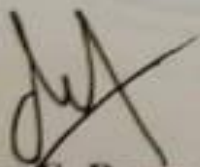
FARUQ NATSIR DJAFAR

T3117023

SKRIPSI

Untuk Memenuhi Salah Satu Syarat Ujian
Guna Memperoleh Gelar Sarjana Program Studi Teknik Informatika
Telah Disetujui Oleh tim Pembimbing
Gorontalo, Juni 2022

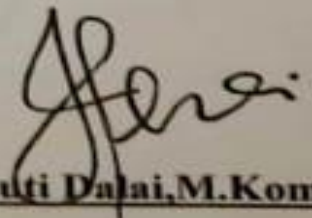
Pembimbing I



Sudirman S. Panna, M.Kom

NIDN. 0924038205

Pembimbing II



Hastuti Dalai, M.Kom

NIDN. 0918038803

PERSETUJUAN SKRIPSI

PENERAPAN ALGORITMA *K-NEAREST NEIGHBOR* (K-NN) UNTUK PREDIKSI LAMA MASA STUDI MAHASISWA DI UNIVERSITAS ICHSAN GORONTALO

Oleh

FARUQ NATSIR DJAFAR

T3117023

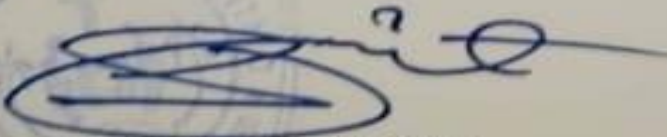
Diperiksa oleh Panitia Ujian Strata Satu (S1)

Universitas Ichsan Gorontalo

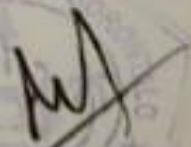
1. Ketua Penguji
Rezqiwati Ishak, M.Kom
2. Anggota
Rofiq Harun, M.Kom
3. Anggota
Maryam Hasan, S.Kom, M.Kom
4. Anggota
Sudirman S. Panna, M.Kom
5. Anggota
Hastuti Dalai, M.Kom

Mengetahui

Dekan Fakultas Ilmu Komputer


Jerry Karim, M.Kom
NIDN. 0918077302

Ketua Program Studi


Sudirman S. Panna, M.Kom
NIDN. 0924038205

PERNYATAAN SKRIPSI

Dengan ini saya menyatakan bahwa :

1. Karya tulis (Skripsi) saya ini adalah asli dan belum pernah diajukan untuk mendapat gelar akademik (Sarjana) baik di Universitas Ichsan Gorontalo maupun di perguruan tinggi lainnya,
2. Karya tulis (Skripsi) saya ini adalah murni gagasan, rumusan, dan penelitian saya sendiri, tanpa bantuan pihak lain, kecuali arahan dari tim Pembimbing.
3. Dalam karya tulis (Skripsi) saya ini tidak terdapat karya atau pendapat yang telah dipublikasikan orang lain, kecuali secara tertulis dicantumkan sebagai acuan/sitasi dalam naskah dan dicantumkan pula dalam daftar Pustaka.
4. Pernyataan ini saya buat dengan sesungguhnya dan apabila dikemudian hari terdapat penyimpangan dan ketidak benaran dalam pernyataan ini, maka saya bersedia menerima sanksi akademik berupa pencabutan gelar yang telah diperoleh karena karya tulis ini, serta sanksi lainnya sesuai dengan norma-norma yang berlaku di Universitas Ichsan Gorontalo.

Gorontalo, Juni 2022

Saya membuat pernyataan ini,



Faruq Natsir Djafar

ABSTRACT

FARUQ NATSIR DJAFAR. T3117023. THE APPLICATION OF THE K- NEAREST NEIGHBOR (K-NN) ALGORITHM FOR THE PREDICTION OF STUDENTS' LENGTH OF STUDY AT UNIVERSITAS ICHSAN GORONTALO

Along with the development of the education field in Indonesia, State Universities and Private Universities compete very tightly in providing their performance in producing the best and quality graduates. This research uses the K-Nearest Neighbor algorithm implemented at Universitas Ichsan Gorontalo. The attributes employed are divided into two, namely the input variable and the output variable. The input variable covers the 1st-semester achievement index, 2nd-semester achievement index, and 3rd-semester achievement index. The output variable covers graduation time. The results of this study indicate that it can predict the students' length of study at Universitas Ichsan Gorontalo using the K-Nearest Neighbor algorithm with an error value of 0.4 for K=5.



Keywords: prediction, length of study, students, KNN

ABSTRAK

FARUQ NATSIR DJAFAR. T3117023. PENERAPAN ALGORITMA *K-NEAREST NEIGHBOR* (K-NN) UNTUK PREDIKSI LAMA MASA STUDI MAHASISWA DI UNIVERSITAS ICHSAN GORONTALO

Seiring dengan perkembangan dunia pendidikan di Indonesia, perguruan tinggi negeri (PTN) maupun perguruan tinggi swasta (PTS) bersaing begitu ketat dalam memberikan performanya mencetak lulusan-lulusan yang terbaik dan berkualitas. Penelitian ini menggunakan algoritma *K-Nearest Neighbor* yang telah dilaksanakan pada Universitas Ichsan Gorontalo, atribut yang digunakan terbagi atas 2 yaitu variabel *input* diantaranya indeks prestasi semester 1, indeks prestasi semester 2 dan indeks prestasi semester 3 juga variabel *output* yaitu waktu lulus. Hasil dari penelitian ini dapat memprediksi waktu lama studi mahasiswa di Universitas Ichsan Gorontalo menggunakan algoritma *K-Nearest Neighbor* yang nilai *error*-nya 0,4 untuk $K=5$.

Kata Kunci : Prediksi, Lama Masa Studi Mahasiswa, *KNN*

KATA PENGANTAR

Bismillahirrahmanirrahim

Puji syukur kepada Allah SWT, atas berkat rahmat dan hidayah-Nya penulis dapat menyelesaikan penyusunan skripsi ini dengan Judul **“Penerapan Algoritma *K-Nearest Neighbor* (K-NN) Untuk Prediksi Lama Masa Studi Mahasiswa Di Universitas Ichsan Gorontalo”** dapat terselesaikan. Sholawat serta salam selalu tercurahkan kepada Nabi besar kita Muhammad SAW yang membawa kita dari alam kebodohan menuju alam yang penuh dengan pengetahuan dan teknologi.

Penulis menyadari sepenuhnya bahwa skripsi ini masih banyak terdapat kekurangan dan kekeliruan. Oleh karena itu penulis mengharapkan saran dan kritikan yang sifatnya membangun demi kesempurnaan skripsi ini. Penulis banyak mendapat bantuan dari berbagai pihak. Oleh karena itu penulis menyampaikan terimakasih yang sebesar-besarnya kepada :

1. Bapak Muhammad Ichsan Gaffar, SE, M.Ak, selaku Ketua Yayasan Pengembangan Ilmu Pengetahuan dan Teknologi (YPIPT) Ichsan Gorontalo;
2. Bapak Dr. Abdul Gaffar La Tjokke, M.Si selaku Rektor Universitas Ichsan Gorontalo;
3. Bapak Jorry Karim, S.Kom, M.Kom, selaku Dekan Fakultas Ilmu Komputer Universitas Ichsan Gorontalo;
4. Bapak Sudirman Melangi, S.Kom, M.Kom, selaku Wakil Dekan I Bidang Akademik & Bidang Kemahasiswaan Fakultas Ilmu Komputer Universitas Ichsan Gorontalo;
5. Ibu Irma Surya Kumala Idris, S.Kom, M.Kom, selaku Pembantu 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. Sekaligus menjadi

Pembimbing Utama yang telah membimbing penulis selama mengerjakan Skripsi ini;

7. Ibu Hastuti Dalai, M.Kom, selaku Pembimbing pendamping yang telah membimbing penulis selama mengerjakan Skripsi ini;
8. Bapak dan Ibu Dosen Universitas Ichsan Gorontalo yang telah mendidik dan mengajarkan berbagai disiplin ilmu kepada penulis;
9. Kedua orang tua saya tercinta (Natsir Djafar dan Jusna Husain S.Pd) telah memberikan kasih sayang, dukungan, doa, dan semangat kepada penulis untuk berjuang dalam menggapai cita-cita;
10. Kepada Mawar Eva de Jongh, yang telah memberikan banyak bantuan, doa dan juga dukungan moril yang sangat besar kepada penulis;
11. Rekan-rekan seperjuangan yang telah banyak memberikan bantuan dan dukungan moril yang sangat besar kepada penulis;
12. Kepada semua pihak yang ikut membantu dalam penyelesaian skripsi ini yang tak sempat penulis sebutkan satu-persatu.

Semoga Allah SWT melimpahkan balasan atas jasa-jasa mereka kepada kami. Penulis menyadari sepenuhnya bahwa apa yang telah dicapai ini masih jauh dari kesempurnaan dan masih terdapat kekurangan. Oleh karena itu, penulis sangat mengharapkan adanya kritik dan saran yang konstruktif. Akhirnya penulis berharap semoga hasil yang telah dicapai ini dapat bermanfaat bagi kita semua, Aamiin.

Gorontalo, Juni 2022

Penulis

DAFTAR ISI

HALAMAN SAMPUL

PENGESAHAN SKRIPSI	i
PERSETUJUAN SKRIPSI	ii
PERNYATAAN SKRIPSI	iii
ABSTRACT	iv
ABSTRAK	v
KATA PENGANTAR	vi
DAFTAR ISI	viii
DAFTAR GAMBAR	x
DAFTAR TABEL	xi
BAB 1 PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Identifikasi Masalah	3
1.3 Rumusan Masalah	3
1.4 Tujuan Penelitian.....	3
1.5 Manfaat Penelitian.....	3
BAB II LANDASAN TEORI	5
2.1 Tinjauan Studi	5
2.2 Tinjauan Pustaka	7
2.2.1 Prediksi.....	7
2.2.2 Kelulusan Studi Mahasiswa	7
2.2.3 Data Mining.....	8
2.2.4 Algoritma <i>K-Nearest Neighbor (K-NN)</i>	10
2.2.5 Konstruksi Sistem	14
2.2.6 <i>Mean Square Error (MSE)</i>	14
2.3 Kerangka Pikir.....	16
BAB III METODE PENELITIAN	17
3.1 Jenis, Metode, Subjek, Objek, Waktu, Dan Lokasi Penelitian.....	17
3.2 Pengumpulan Data	17

3.3	Pemodelan Abstraksi	19
3.3.1	Pengembangan Model	19
3.3.2	Evaluasi Model	20
BAB IV	HASIL PENELITIAN	21
4.1	Hasil Pengumpulan Data	21
4.1.1	Pra Pengolahan Data	21
4.2	Hasil Pemodelan	23
4.2.1	Penerapan Algoritma <i>K-Nearest Neighbor (K-NN)</i>	23
4.2.2	Evaluasi Model Dengan <i>Mean Square Error (MSE)</i>	25
BAB V	Pembahasan Hasil	28
5.1	Objek dan Lokasi Penelitian	28
5.1.1	Profil Universitas Ichsan Gorontalo	28
5.1.2	Sejarah Singkat Universitas Ichsan Gorontalo	29
5.2	Pembahasan Hasil Eksperimen	29
5.2.1	Pembahasan Pengumpulan Data	30
5.2.2	Pembahasan Pra Pengolahan Data	30
5.2.3	Hasil Penerapan Algoritma <i>K-Nearest Neighbor (K-NN)</i>	31
5.2.4	Pembahasan Hasil <i>Mean Square Error (MSE)</i>	31
BAB VI	Kesimpulan dan Saran	32
6.1	Kesimpulan	32
6.2	Saran	32
DAFTAR PUSTAKA		
LAMPIRAN-LAMPIRAN		

DAFTAR GAMBAR

Gambar 2. 1 <i>Knowledge Discovery Database (KDD)</i>	9
Gambar 2. 2 Kerangka Pikir	16
Gambar 3. 1 Pemodelan.....	19
Gambar 4.1 <i>Outlier</i> Indeks Prestasi Semester 1	22
Gambar 4.2 <i>Outlier</i> Indeks Prestasi Semester 2.....	22
Gambar 4.3 <i>Outlier</i> Masa Studi.....	23

DAFTAR TABEL

Tabel 2. 1 Penelitian Tentang Prediksi <i>K-Nearest Neighbor</i>	5
Tabel 2. 2 Data <i>Training</i>	11
Tabel 2. 3 Perhitungan Jarak menggunakan <i>Euclidean Distance</i>	13
Tabel 2. 4 Jarak terdekat sebanyak $k = 5$	14
Tabel 4.1 Hasil Pengumpulan Data.....	21
Tabel 4.2 Data <i>Training</i>	23
Tabel 4.3 Perhitungan Jarak Menggunakan <i>Euclidean Distance</i>	25
Tabel 4.4 Jarak Terdekat Sebanyak $K=5$	25
Table 4.5 Evaluasi <i>Mean Square Error</i>	27

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Lulus tepat waktu merupakan salah satu indikator keberhasilan suatu perguruan tinggi dalam memberikan layanan pendidikan bagi masyarakat. Di Negara Indonesia baik perguruan tinggi negeri (PTN) maupun perguruan tinggi swasta (PTS) tentunya semua ingin berlomba untuk mendapatkan predikat baik di mata masyarakat dalam rangka meningkatkan animo masyarakat terhadap perguruan tinggi tersebut yang salah satunya adalah dengan memenuhi indikator jumlah persentase mahasiswa lulus tepat waktu yang tinggi. Selain untuk meningkatkan animo masyarakat terhadap perguruan tinggi, lama masa studi juga merupakan salah satu indikator penilaian dalam pemeringkatan akreditasi perguruan tinggi.

Universitas Ichsan Gorontalo merupakan salah satu perguruan tinggi swasta yang beroperasi di wilayah provinsi Gorontalo dengan melayani sejumlah masyarakat yang berasal dari wilayah provinsi Gorontalo dan beberapa provinsi sekitar seperti Sulawesi Utara, Sulawesi Tengah, dan beberapa provinsi di wilayah bagian Timur Indonesia. Universitas Ichsan Gorontalo saat ini tercatat memiliki mahasiswa yang sedang menempuh pendidikan sebesar kurang lebih 5506 mahasiswa dan memiliki 219 Dosen Tetap, dengan jumlah program studi sebanyak 11 program studi. Sejak berdiri hingga saat ini tercatat Universitas Ichsan Gorontalo telah menghasilkan lulusan sebanyak kurang lebih 10.300 lulusan.

Jumlah lulusan Universitas Ichsan Gorontalo yang sebesar kurang lebih 10.300 lulusan tersebut, ditemukan bahwa hanya sekitar 45% mahasiswa yang dapat menyelesaikan masa studinya secara tepat waktu, tentunya hal tersebut menjadi indikator yang kurang baik bagi perguruan tinggi tersebut dalam menyelenggarakan proses pendidikan yang tentunya akan berakibat pada menurunnya tingkat kepercayaan masyarakat pada Universitas Ichsan Gorontalo

dalam hal pelaksanaan pendidikan. Sehingga Universitas Ichsan Gorontalo membutuhkan sebuah solusi yang memadai untuk mengatasi masalah tersebut guna perbaikan kedepannya seperti pemanfaatan *Artificial Intelligent* untuk memprediksi lama masa studi mahasiswa. Solusi dengan pendekatan data science yang digunakan akan melakukan proses analitik menggunakan algoritma *K-Nearest Neighbor* terhadap data pencapaian akademik mahasiswa yang bersumber dari hasil studi mahasiswa pada beberapa semester awal.

Menurut Pandi algoritma *K-Nearest Neighbors* (KNN) sebuah data baru diklasifikasikan berdasarkan jarak data baru tersebut dengan tingkat kemiripan data baru terdekat terhadap data pola. Jumlah data tetangga terdekat ditentukan dan dinyatakan dengan k . Penentuan nilai k terbaik dapat ditentukan dengan optimasi parameter, misalnya dengan menggunakan *k-Fold Cross Validation* yang merupakan salah satu metode yang digunakan untuk mengetahui rata-rata keberhasilan dari suatu sistem dengan cara melakukan perulangan dengan mengacak atribut masukan sehingga sistem tersebut teruji untuk beberapa atribut input yang acak. Nilai $k=1$ pada algoritma *K-Nearest Neighbors* (KNN) bukan satu-satunya faktor yang menentukan pemilihan *k-Optimal*, karena pada $k=1$ perhitungan jarak dan penentuan hasil prediksi dipengaruhi oleh perhitungan data. Untuk memutuskan satu nilai k terbaik atau *k-Optimal* maka dilihat dari tingkat akurasi pada hasil pencarian menggunakan metode *k-Fold Cross Validation*. Nilai *k-Optimal* pada algoritma *K-Nearest Neighbor* untuk prediksi kelulusan tepat waktu mahasiswa berdasarkan IP sampai dengan semester 4 adalah $k=5$. Dari proses *k-Fold Cross Validation* didapatkan tingkat akurasi untuk $k=5$ pada algoritma *K-Nearest Neighbor* untuk prediksi kelulusan tepat waktu mahasiswa berdasarkan IP sampai dengan semester 4 adalah sebesar 80,00%. [1]

Dengan memberikan informasi mahasiswa yang berpotensi mengalami keterlambatan dalam penyelesaian studi, diharapkan dapat memberikan masukan kepada perguruan tinggi Universitas Ichsan Gorontalo dalam merancang kebijakan yang tepat dalam menangani masalah-masalah yang

berkaitan dengan lama masa studi mahasiswa agar nilai presentase mahasiswa lulus tepat waktu dapat lebih meningkat.

Berdasarkan uraian di atas, pendekatan *Artificial Intelligence* (AI) khususnya di bidang data *science* dapat menjadi salah satu solusi yang memungkinkan untuk digunakan dalam menyelesaikan permasalahan yang di hadapi oleh Universitas Ichsan Gorontalo maka dari itu penulis mengusulkan skripsi dengan judul **“Penerapan Algoritma *K-Nearest Neighbor* Untuk Prediksi Lama Masa Studi Mahasiswa Di Universitas Ichsan Gorontalo”**.

1.2 Identifikasi Masalah

Berdasarkan latar belakang yang telah diuraikan sebelumnya, permasalahan yang dihadapi oleh Universitas Ichsan Gorontalo adalah ditemukan sekitar 45 % mahasiswa mengalami keterlambatan dalam menyelesaikan studinya

1.3 Rumusan Masalah

Berdasarkan identifikasi masalah di atas, maka peneliti merumuskan masalah yang nantinya akan dikembangkan dalam hasil prediksi ini, yaitu bagaimana hasil penerapan metode *K-Nearest Neighbor* dalam memprediksi lama masa studi mahasiswa di Universitas Ichsan Gorontalo.

1.4 Tujuan Penelitian

Sesuai dengan permasalahan di atas, maka tujuan dari penelitian ini adalah mengetahui hasil penerapan metode *K-Nearest Neighbor* dalam memprediksi lama masa studi mahasiswa di Universitas Ichsan Gorontalo.

1.5 Manfaat Penelitian

Penelitian ini dapat memberikan manfaat sebagai berikut:

1. Manfaat Teoritis

Hasil penelitian ini dapat memberikan masukan bagi perkembangan ilmu pengetahuan dan teknologi, khususnya pada bidang ilmu Komputer berupa pemuktahiran metode *K-Nearest Neighbor*.

2. Manfaat Praktis

Sumbangan pemikiran, karya, bahan pertimbangan, atau solusi bagi *software developer* guna mendukung pengambilan keputusan dalam rangka menghasilkan *software* yang berkualitas sehingga dapat mengatasi masalah mahasiswa yang berpotensi untuk lulus tidak tepat waktu agar jumlah lulusan tepat waktu di perguruan tinggi tersebut dapat meningkat.

BAB II

LANDASAN TEORI

2.1 Tinjauan Studi

Penelitian tentang prediksi lama masa studi mahasiswa menggunakan metode *K-Nearest Neighbor* telah dilakukan oleh beberapa peneliti sebelumnya dapat dilihat di tabel berikut ini :

Tabel 2. 1 Penelitian Tentang Prediksi *K-Nearest Neighbor*

NO.	PENELITI	JUDUL/TAHUN	METODE	HASIL
1.	Mustakim1, Giantika Oktaviani F	Algoritma <i>K-Nearest Neighbor Classification</i> Sebagai Sistem Prediksi Predikat Prestasi Mahasiswa./2016	<i>K-NEAREST NEIGHBOR</i>	Perhitungan algoritma K-NN yang diterapkan dalam memprediksi predikat prestasi Mahasiswa mampu menghasilkan akurasi dengan nilai 82%.Pengujian algoritma ini dilakukan menggunakan perhitungan <i>confusion matriks</i> yaitu membandingkan predikat pada semester sebelumnya dengan predikat hasil prediksi[2].
2.	Eri Sasmita Susanto, Kusrini & Hanif Al Fatta	PREDIKSI KELULUSAN MAHASISWA MAGISTER TEKNIK INFORMATIKA UNIVERSITAS AMIKOM YOGYAKARTA	<i>K-NEAREST NEIGHBOR</i>	Dari hasil pengujian prediksi kelulusan mahasiswa dengan metode <i>K-Nearest Neighbor</i> berdasarkan nilai 4 semester menggunakan

NO.	PENELITI	JUDUL/TAHUN	METODE	HASIL
		MENGGUNAKAN METODE <i>K-NEAREST NEIGHBOR</i> ./2018		Rapidminer didapatkan kesimpulan sebagai berikut ini: 1. <i>K-optimal</i> dan <i>k-fold cross validation</i> dalam prediksi kelulusan mahasiswa dengan metode <i>K-Nearest Neighbor</i> berdasarkan nilai 4 semester adalah <i>k=14</i> dan <i>k-fold=5</i> . 2. Berdasarkan penerapan <i>k=14</i> dan <i>k-fold=5</i> menghasilkan performa yang terbaik dalam memprediksi kelulusan mahasiswa dengan metode <i>K-Nearest Neighbor</i> menggunakan indeks prestasi 4 semester dengan nilai akurasi= 98,46%, precision= 99.53% dan recall =97.64%. [3]
3.	Mutiara, Irwan dan Andi	Penerapan <i>K-Optimal</i> Pada Algoritma Knn Untuk Prediksi Kelulusan Tepat Waktu Mahasiswa Program Studi Ilmu Komputer Fmipa Unlam Berdasarkan Ip Sampai Dengan Semester 4./2015	<i>K-NEAREST NEIGHBOR</i>	Dari proses <i>k-Fold Cross Validation</i> didapatkan tingkat akurasi untuk <i>k=5</i> pada algoritma kNN untuk prediksi kelulusan tepat waktu mahasiswa berdasarkan IP

NO.	PENELITI	JUDUL/TAHUN	METODE	HASIL
				sampai dengan semester 4 adalah sebesar 80,00%. [1]

2.2 Tinjauan Pustaka

2.2.1 Prediksi

Prediksi adalah suatu proses memperkirakan secara sistematis tentang sesuatu yang paling mungkin terjadi di masa depan dengan berdasarkan informasi masa lalu dan sekarang yang dimiliki, agar kesalahan (selisih antara sesuatu yang terjadi dengan hasil perkiraan) dapat diperkecil. Prediksi tidak harus memberikan jawaban secara pasti kejadian yang akan terjadi, melainkan berusaha untuk mencari jawaban sedekat mungkin yang akan terjadi. [4]

2.2.2 Kelulusan Studi Mahasiswa

Kriteria kelulusan mahasiswa dari suatu program studi dirumuskan dalam bentuk Standar Kompetensi Lulusan yang terdapat dalam rancangan kurikulum. Secara khusus, Pasal 1 butir 4 Peraturan Pemerintah No. 19 tahun 2005 tentang Standar Nasional Pendidikan, menyebutkan bahwa “Standar Kompetensi Lulusan” adalah kualifikasi kemampuan lulusan yang mencakup sikap, pengetahuan, dan keterampilan.

Pada suatu perguruan tinggi contohnya Universitas Ichsan Gorontalo, jumlah mahasiswa yang lulus tidak tepat waktu tergolong lebih banyak dari pada mahasiswa yang lulus tepat waktu. Mahasiswa dapat dikatakan lulus tepat waktu jika mampu menyelesaikan masa studinya tidak lebih dari 4 tahun, sedangkan mahasiswa yang menyelesaikan masa studinya lebih dari 4 tahun termasuk dalam mahasiswa yang lulus tidak tepat waktu.

Ada beberapa faktor yang mempengaruhi lamanya jangka waktu masa studi mahasiswa, termasuk elemen luar dan dalam. Faktor internal yang menyebabkan lamanya masa studi mahasiswa antara lain; (1) tidak adanya dorongan hati, dan (2) memilih jurusan yang salah. Sedangkan faktor eksternal yang menyebabkan lamanya masa studi mahasiswa antara lain; (1) terlalu santai karena jauh dari

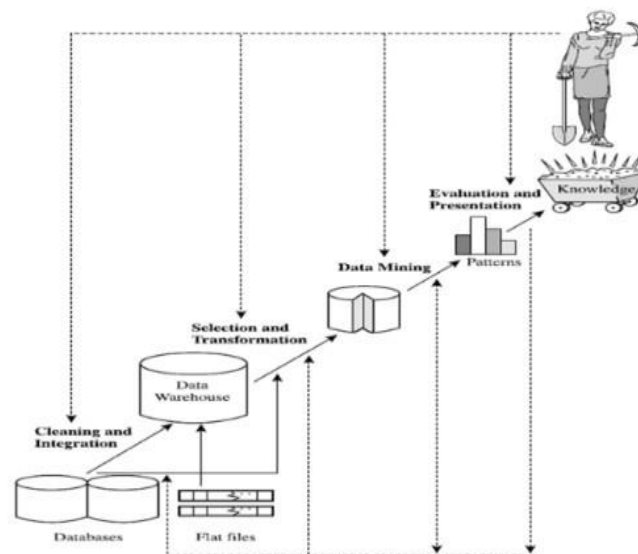
pengawasan orang tuanya, (2) terlalu aktif dalam mengikuti perkumpulan mahasiswa, (3) berlebihan dalam menjalankan hobi, (4) bekerja, dan (5) tidak ada jaminan menemukan pekerjaan baru setelah lulus [5].

Menurut Peraturan akademik Universitas Ichsan Gorontalo tahun 2019, BAB V, Pasal 24, mahasiswa dinyatakan lulus program sarjana apabila [6]:

- 1) Mahasiswa dinyatakan lulus apabila telah mengumpulkan jumlah minimal SKS yang disyaratkan untuk program studi bersangkutan, IPK minimal 2,50, tidak ada nilai E, nilai D paling banyak 10 % dari jumlah SKS yang ditempuh, dan telah lulus ujian akhir;
- 2) Mahasiswa yang telah lulus diwisuda sesuai dengan ketentuan yang berlaku di Universitas Ichsan Gorontalo;
- 3) Mahasiswa yang diyudisium berhak memakai gelar akademik sesuai dengan ketentuan yang berlaku; dan
- 4) Persyaratan wisuda wajib dipenuhi dan prosesi wisuda wajib diikuti oleh mahasiswa yang lulus, kecuali terdapat alasan yang dapat diterima.
- 5) Mahasiswa yang berhalangan atau tidak menghadiri prosesi wisuda sebagaimana pada ayat (4) tidak membatalkan gelar kesarjanaannya

2.2.3 Data Mining

Data *mining* merupakan proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar [5].



Gambar 2. 1 *Knowledge Discovery Database (KDD)*

Proses KDD sebagai berikut:

1. Pembersihan data (*Data Cleaning*).

Proses pembersihan data atau data cleaning dilakukan untuk menghilangkan noise dan data yang tidak konsisten.

2. Integrasi data (*Data Integration*).

Proses data integrasi adalah proses menggabungkan data dari sumber data yang berbeda.

3. Seleksi data (*Data Selection*).

Seleksi data adalah proses memilih data atau atribut yang relevan untuk atribut ini. Pada tahap ini dilakukan analisis korelasi atribut data. Atribut – atribut data tersebut dicek apakah relevan untuk dilakukan penambangan data.

4. Transformasi data (*Data Transformation*).

Transformasi atau data transformation proses menggabungkan data ke dalam bank yang sesuai untuk ditambang.

5. Penambangan data (*Data Mining*)

Langkah ini adalah langkah paling penting yaitu melakukan pengaplikasian metode yang tepat untuk pola data.

6. Evaluasi pola (*Pattern Evaluation*).

Pada langkah ini dilakukan identifikasi pola dalam bentuk pengetahuan berdasarkan beberapa pengukuran yang penting.

7. Presentasi pengetahuan (*Knowledge Presentation*).

Pada langkah ini dilakukan proses penyajian pengetahuan dari hasil penambahan data. hasil klasifikasi data nasabah akan ditampilkan ke dalam bentuk yang mudah dipahami *user* pengguna.

2.2.4 Algoritma *K-Nearest Neighbor* (*K-NN*)

K-Nearest Neighbors (*K-NN*) adalah suatu metode yang menggunakan algoritma supervised dimana hasil dari *query instance* yang baru diklasifikasikan berdasarkan mayoritas dari label class pada *K-NN*. Tujuan dari algoritma *K-NN* adalah mengklasifikasikan objek baru berdasarkan atribut dan training data.

Algoritma *K-NN* bekerja berdasarkan jarak terdekat dari *query instance* ke training data untuk menentukan *KNN*-nya. Salah satu cara untuk menghitung jarak dekat atau jauhnya tetangga menggunakan metode *euclidean distance*.

Euclidean Distance sering digunakan untuk menghitung jarak. *Euclidean distance* berfungsi menguji ukuran yang bisa digunakan sebagai interpretasi kedekatan jarak antara dua obyek, di bawah ini merupakan rumus *euclidean distance* [7]:

$$\left(\sum_{k=1}^m (x_{ik} - x_{jk})^2 \right)^{1/2}$$

Dimana :

x_{ik} = nilai X pada training data

x_{jk} = nilai X pada testing data

m = batas jumlah banyaknya data

Jika hasil nilai dari rumus di atas besar maka akan semakin jauh tingkat keserupaan antara kedua objek dan sebaliknya jika hasil nilainya semakin kecil maka akan semakin dekat tingkat keserupaan antar objek tersebut. Objek yang dimaksud adalah *training* data dan *testing* data.

Dalam perhitungan ini, nilai terbaik dari k bergantung pada seberapa banyak informasi. Ukuran nilai k yang besar sebenarnya bukan nilai k terbaik begitu juga sebaliknya.

Langkah-langkah untuk menghitung algoritma K -NN:

1. Tentukan nilai k .
2. Hitung kuadrat *Euclidean distance* (*instance query*) dari setiap objek ke data pelatihan yang diberikan.
3. Kemudian mengurutkan obyek-obyek tersebut ke dalam kelompok-kelompok yang memiliki *Euclidean distance* terkecil.
4. Mengumpulkan label kelas Y (*K-Nearest Neighbor*).
5. Dengan menggunakan kategori *Nearest Neighbor* yang merupakan mayoritas, nilai *instance query* yang dihitung dapat diprediksi.

Berikut adalah contoh perhitungan dengan metode K -NN. dengan menerima informasi persiapan sebanyak 20 informasi dari angkatan lulus [8]:

Tabel 2. 2 Data Training

NO	NIM	IPS 1	IPS 2	IPS 3	IPS 4	TOTAL SKS	STATUS
1	135610075	3,19	3,52	3,50	3,13	91	LC
2	135610082	3,19	3,52	3,33	3,22	91	LC
3	135610123	3,86	3,65	3,36	3,68	88	LC
4	135610140	3,38	3,38	3,29	3,33	93	LC
5	135610152	2,62	2,86	3,60	3,42	86	LC
6	135610157	3,62	3,38	3,26	3,67	89	LC
7	085610034	1,00	1,38	0,00	1,13	70	LL
8	095610111	2,14	0,79	1,62	0,13	68	LL
9	095610116	1,10	1,55	0,00	2,20	63	LL
10	105610080	1,48	2,62	2,29	1,54	70	LL
11	105610087	2,24	0,00	1,92	2,57	74	LL
12	115610022	1,76	1,93	2,43	1,90	70	LL
13	115610062	2,05	2,19	1,59	1,15	82	LL
14	115610118	2,14	2,50	2,80	2,33	82	LL
15	125610042	2,10	2,24	2,50	2,52	81	LT
16	125610050	2,14	2,56	2,90	2,67	83	LT
17	125610055	1,56	2,19	2,19	2,00	71	LT
18	125610056	1,86	1,94	2,39	1,72	73	LT
19	125610069	2,33	2,67	2,35	1,26	81	LT
20	125610083	2,71	2,05	2,61	2,50	80	LT

Keterangan Status:

Lulus Cepat (LC)	: 7 semester
Lulus Tepat (LT)	: 8 sampai dengan 10 semester
Lulus Lambat (LL)	: lebih dari 10 semester

Terdapat data baru mahasiswa sebagai data uji kemudian dilakukan perhitungan berdasarkan metode $K-NN$ dengan *euclidean distance* sebagai berikut :

Nim	: 165610122
IPS 1	: 3,83
IPS 2	: 3,73
IPS 3	: 3,50
IPS 4	: 3,10
Total sks semester 4	: 77

Dari data baru untuk mahasiswa dengan nim 165610122 tersebut dihitung jarak dengan menggunakan rumus (2.2) nilai K adalah 5 sebagai berikut :

1. Nilai $d(x_1-y_1)^2 =$

$$\begin{aligned}
 & \sqrt{(3,83 - 3,19)^2 + (3,73 - 3,52)^2 + (3,50 - 3,50)^2 + (3,10 - 3,13)^2 + (77 - 91)^2} \\
 &= \sqrt{0,4096 + 0,0441 + 0 + 0,0009 + 196} \\
 &= \sqrt{196,4546} \\
 &= 14,0162
 \end{aligned}$$

2. Nilai $d(x_1-y_2)^2 =$

$$\begin{aligned}
 & \sqrt{(3,83 - 3,19)^2 + (3,73 - 3,52)^2 + (3,50 - 3,33)^2 + (3,10 - 3,22)^2 + (77 - 91)^2} \\
 &= \sqrt{0,4096 + 0,0441 + 0,0289 + 0,0144 + 196} \\
 &= \sqrt{196,4970}
 \end{aligned}$$

$$=14,0177$$

3. Nilai $d(x_1-y_3)^2$

$$\sqrt{(3,83 - 3,86)^2 + (3,73 - 3,65)^2 + (3,50 - 3,36)^2 + (3,10 - 3,38)^2 + (77 - 88)^2}$$

$$=\sqrt{0,0009 + 0,0064 + 0,0196 + 0,3364 + 121}$$

$$=\sqrt{121,3633}$$

$$=11,0165$$

4. Nilai $d(x_1-y_4)^2$

$$\sqrt{(3,83 - 3,38)^2 + (3,73 - 3,38)^2 + (3,50 - 3,29)^2 + (3,10 - 3,33)^2 + (77 - 93)^2}$$

$$= \sqrt{0,2025 + 0,12225 + 0,0441 + 0,0529 + 256}$$

$$=\sqrt{256,4220}$$

$$=16,0132$$

Berikut ini adalah tabel hasil perhitungan jarak menggunakan *Euclidean Distance*:

Tabel 2. 3 Perhitungan Jarak menggunakan *Euclidean Distance*

NO	NIM	JARAK EUCLIDEAN	RANKING
1	135610075	14,0162	17
2	135610082	14,0177	18
3	135610123	11,0165	15
4	135610140	16,0132	20
5	135610152	9,1287	13
6	135610157	12,0229	16
7	085610034	8,8692	12
8	095610111	10,2399	14
9	095610116	14,8750	19
10	105610080	7,7235	11
11	105610087	5,3121	5
12	115610022	7,6883	10
13	115610062	6,1637	7

NO	NIM	JARAK EUCLIDEAN	RANKING
14	115610118	5,5183	6
15	125610042	4,7486	2
16	125610050	6,3851	8
17	125610055	6,8155	9
18	125610056	5,1207	4
19	125610069	4,9073	3
20	125610083	3,7721	1

- *Nearest neighbor* ditentukan pada awal adalah $k = 5$, yaitu 5 jarak yang paling kecil :

Tabel 2. 4 Jarak terdekat sebanyak $k = 5$

No	NIM	Jarak	Rangking	Status
1	125610083	3,7721	1	LT
2	125610042	4,7486	2	LT
3	125610069	4,9073	3	LT
4	125610056	5,1207	4	LT
5	105610087	5,3121	5	LL

- Menghitung jumlah status yang lebih banyak muncul. Pada bagian ini, kemunculan status terbanyak adalah LT sebanyak 4 kali, sedangkan kemunculan status LL sebanyak satu kali.
- Kesimpulan dari hasil perkiraan masa studi mahasiswa dengan NIM 165610122 adalah Lulus Tepat Waktu (LT).

2.2.5 Konstruksi Sistem

Konstruksi sistem adalah pengembangan dan pengujian komponen system untuk membuat dan menguji sistem fungsional yang memenuhi persyaratan desain untuk mengimplementasikan antarmuka sistem baru dan sistem produksi yang telah ada.

2.2.6 Mean Square Error (MSE)

Mean Square Error (MSE) adalah rata-rata dari kesalahan *forecast* dikuadratkan. *Mean Square Error (MSE)* menggunakan nilai kuadrat untuk setiap selisih perhitungan yang terjadi. Perbedaannya dengan *Mean Absolute Deviation (MAD)* adalah MSE menilai kesalahan untuk penyimpangan yang lebih ekstrem daripada MAD[9].

Nilai MSE didapat dari persamaan :

$$\text{MSE} = \frac{\sum_{t=1}^n (y_t - y'_t)^2}{n}$$

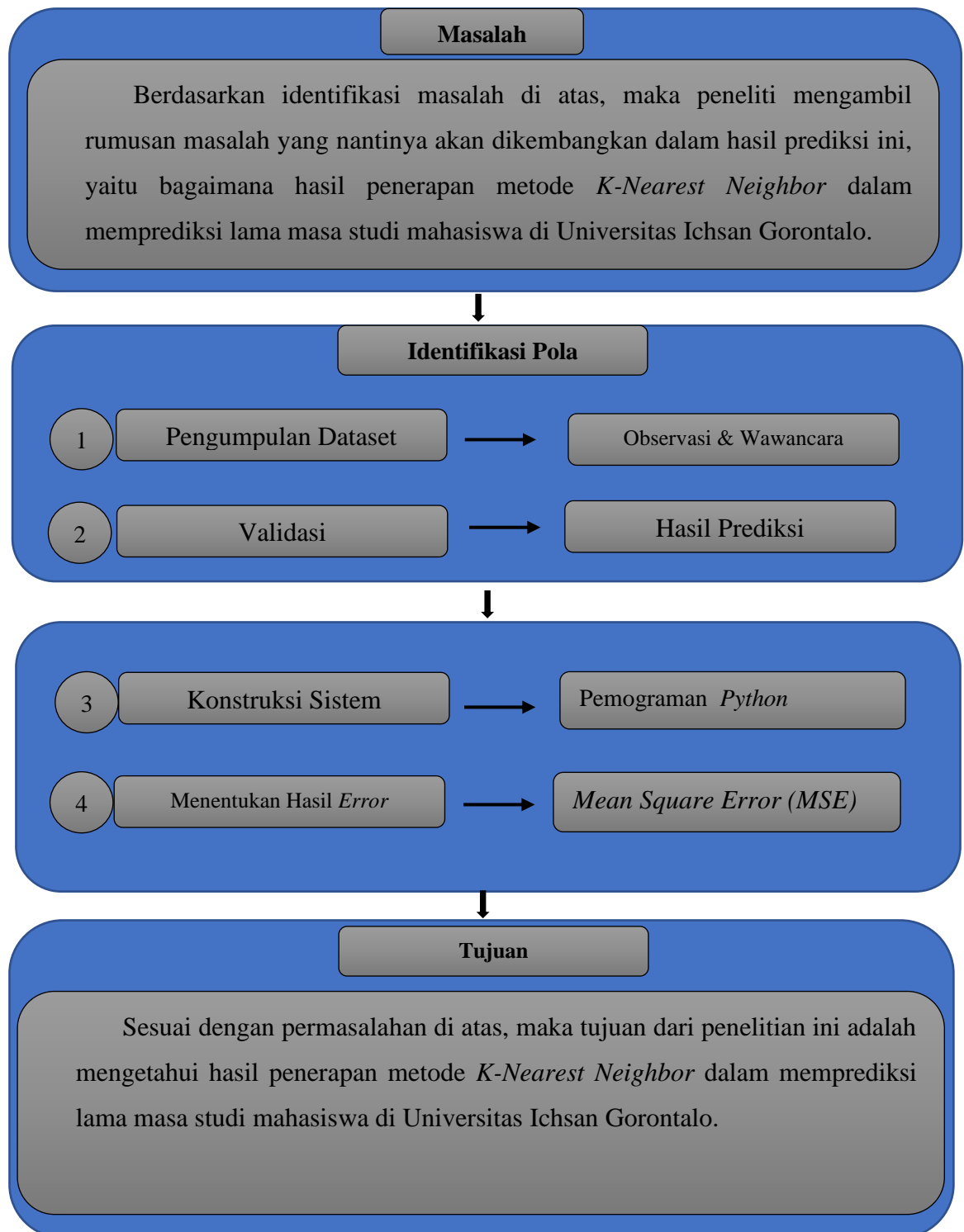
Keterangan:

y_t = data aktual pada periode t

y'_t = data peramalan pada periode t

n = jumlah data

2.3 Kerangka Pikir



Gambar 2. 2 Kerangka Pikir

BAB III

METODE PENELITIAN

3.1 Jenis, Metode, Subjek, Objek, Waktu, Dan Lokasi Penelitian

Berdasarkan dari tingkat penerapan, maka penelitian ini merupakan penelitian terapan yang dipandang dari jenis informasi yang diolah, maka riset ini bagian dari kuantitatif. Begitu pun jika ditinjau dari perilaku terhadap data, maka penelitian ini merupakan penelitian konfirmasi.

Riset ini berpedoman pada aturan *K-Nearest Neighbors* yang berfokus di Universitas Ichsan Gorontalo, dikarenakan sifatnya yang menggambarkan kejadian yang ada pada pabrik maka dikatakan sebagai riset deskriptif. Adapun topik dari riset ini yakni prediksi lama masa studi mahasiswa. Penelitian ini dimulai dari bulan November – Desember 2021 yang berlokasi di Universitas Ichsan Gorontalo JL. Drs. Achmad Nadjamudin No 10, Kota Gorontalo.

3.2 Pengumpulan Data

Untuk mengumpulkan data yang digunakan ada 2 (dua) jenis data yaitu: Data Primer dan Data Sekunder. Data primer merupakan data yang berasal dari penelitian lapangan dan data sekunder berasal dari penelitian kepustakaan

1. Penelitian Data Primer (Lapangan)

Untuk memperoleh data primer yang merupakan data langsung dari objek penelitian yaitu Universitas Ichsan Gorontalo, maka dilakukan dengan teknik:

- 1) Observasi, metode ini memungkinkan analisis sistem mengamati atau meninjau langsung. adapun pada penelitian ini dilakukan dengan cara mengumpulkan data mahasiswa di Universitas Ichsan Gorontalo.
- 2) Wawancara, metode ini digunakan dengan mengajukan beberapa pertanyaan tentang lama masa studi mahasiswa dari setiap fakultas.

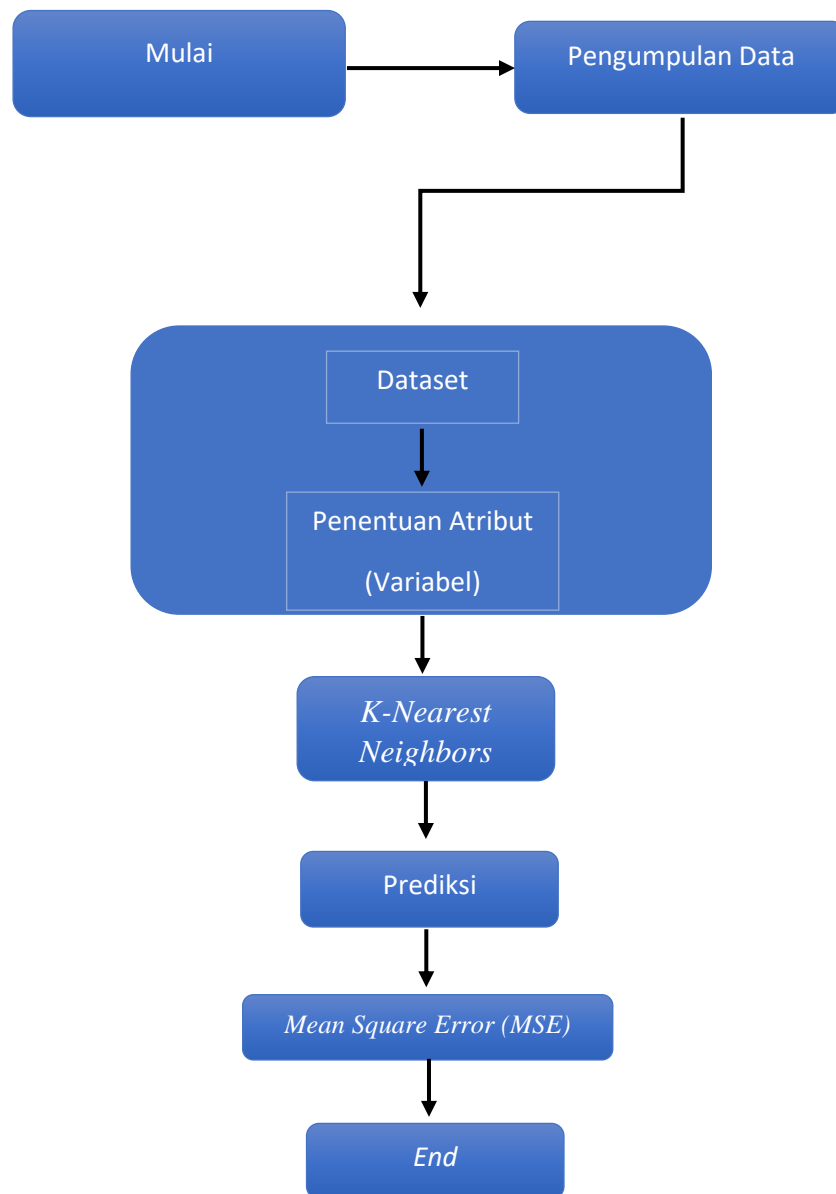
2. Penelitian Data Sekunder (Kepustakaan)

Metode kepustakaan diperlukan untuk mendapatkan data sekunder dengan tujuan melengkapi data primer. Data sekunder didapatkan dari pengkajian kepustakaan yang berisi tentang dasar teori-teori. Metode kepustakaan ini

digunakan oleh analisis sistem dengan cara mengambil contoh dokumen-dokumen yang berhubungan dengan materi penelitian, selain itu juga analisis sistem mencari data mengenai hal-hal atau parameter yang berupa catatan, buku, majalah, dan sebagainya yang berhubungan dengan penelitian ini

Adapun variable data antara lain Indeks Prestasi Semester 1, Indeks Prestasi Semester 2, Indeks Prestasi Semester 3 dan selisih tahun masuk dan tahun keluar (lama masa studi).

3.3 Pemodelan Abstraksi



Gambar 3. 1 Pemodelan

3.3.1 Pengembangan Model

Prosedur atau langkah-langkah pokok prediksi lama masa studi mahasiswa menggunakan metode *K-Nearest Neighbor* dengan alat bantu Pemrograman *Python*, serta *Mean Square Error (MSE)* untuk menguji kinerja sistemnya.

3.3.2 Evaluasi Model

Proses evaluasi bertujuan untuk mengetahui hasil performa dari metode yang digunakan, mengevaluasi semua data uji, kemudian memetakan target keluaran yang dihasilkan ke dalam *Mean Square Error (MSE)* untuk mengevaluasi nilai *errornya*.

BAB IV

HASIL PENELITIAN

4.1 Hasil Pengumpulan Data

Berdasarkan hasil pengumpulan data yang diperoleh dari tempat penelitian yaitu BAAK Universitas Ichsan Gorontalo berjumlah 10397 data (alumni).

Tabel 4. 1 Hasil Pengumpulan Data

No.	Masuk	Lulus	L/P	Prodi	IPS_1	IPS_2	IPS_3	Selisih
2291	01-Sep-22	09-Apr-19	P	S1	3,48	3,65	3,38	5 Thn
2951	25-Ags-08	13-Apr-13	P	S1	3,13	3,14	3,00	5 Thn
3201	07-Sep-09	08-Apr-14	L	S1	3,24	3,08	2,83	5 Thn
4531	24-Ags-17	05 - Jun-21	P	S1	3,76	3,63	3,57	4 Thn
4532	24-Ags-17	05 - Jun-21	P	S1	3,76	4,00	3,74	4 Thn
8031	03-Sep-12	24-Nov-17	P	S1	3,30	2,67	2,79	5,5 Thn
8421	03-Sep-07	27-Nov-11	P	S1	3,00	3,31	3,00	4,5 Thn
8540	25-Ags-15	02-Des-14	P	S1	3,27	2,84	3,60	4,5 Thn
10311	24-Ags-15	10-Des-19	P	S1	3,61	3,42	3,42	4,5 Thn
.....
10381	28-Ags-17	07 - Jun-17	L	S1	3,20	3,67	3,92	4 Thn

4.1.1 Pra Pengolahan Data

1. *Missing Value* (data yang hilang)

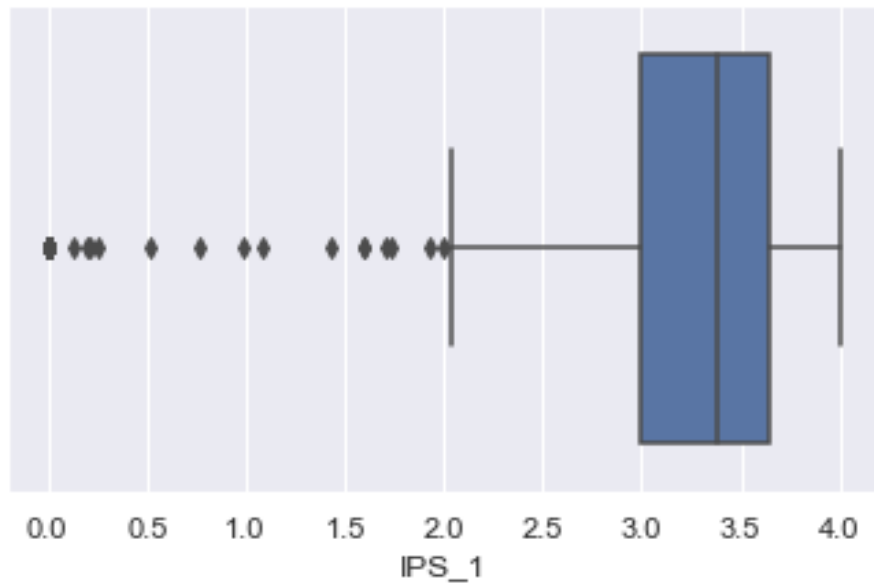
Dari data yang ada, terdapat beberapa *missing value* (data yang hilang), yaitu :

- a. Indeks Prestasi Semester 1 :46
- b. Indeks Prestasi Semester 2 :160
- c. Indeks Prestasi Semester 3 :312
- d. Masa Studi :991

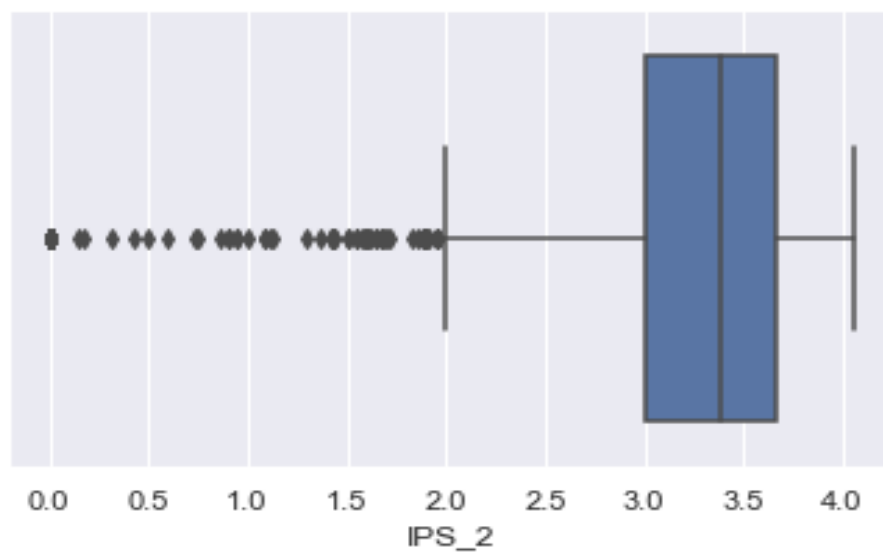
Dari jumlah *missing value* tersebut dilakukan *imputation* menggunakan *Regresi Linear* untuk mengolah *missing value* yang ada.

2. *Outlier* (nilai tidak wajar)

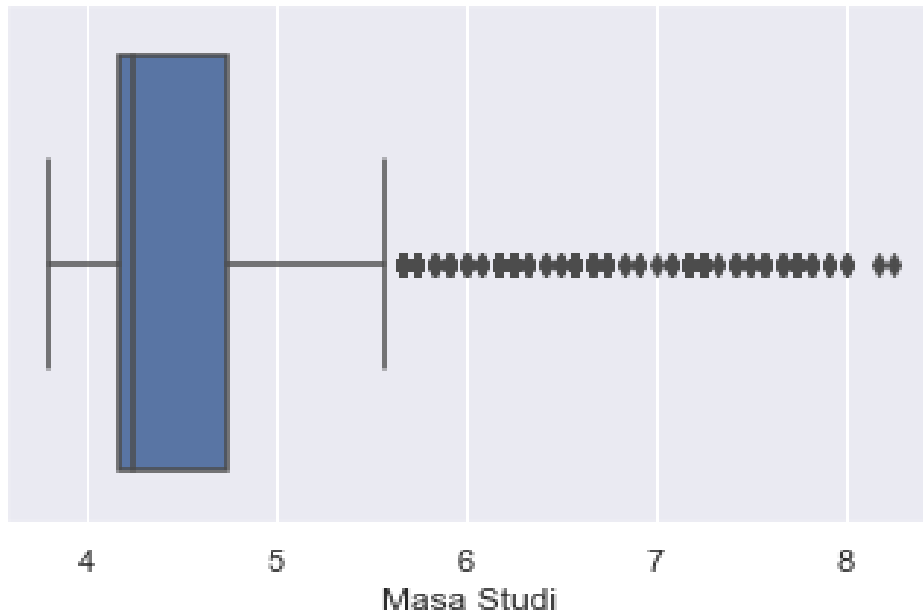
Outlier dapat mengakibatkan masalah serius dalam analisis data. Maka dari itu di butuhkan *InterQuartile Range (IQR)* untuk menarik kesimpulan sekumpulan data dan menyertakan data paling luar.



Gambar 4. 1 *Outlier* Indeks Prestasi Semester 1



Gambar 4. 2 *Outlier* Indeks Prestasi Semester 2



Gambar 4. 3 *Outlier Masa Studi*

4.2 Hasil Pemodelan

4.2.1 Penerapan Algoritma *K-Nearest Neighbor (K-NN)*

Berikut ini merupakan perhitungan menggunakan algoritma *K-NN* dengan mengambil data *training* sebanyak 10 data dari Alumni :

Tabel 4. 2 *Data Training*

No.	IPS_1	IPS_2	IPS_3	Selisih
2291	3,48	3,65	3,38	5 Tahun
2951	3,13	3,14	3,00	5 Tahun
3201	3,24	3,08	2,83	5 Tahun
4531	3,76	3,63	3,57	4 Tahun
4532	3,76	4,00	3,74	4 Tahun
8031	3,30	2,67	2,79	5,5 Tahun
8421	3,00	3,31	3,00	4,5 Tahun
8540	3,27	2,84	3,60	4,5 Tahun
10311	3,61	3,42	3,42	4,5 Tahun
10381	3,20	3,67	3,92	4 Tahun

Terdapat data baru mahasiswa sebagai data uji kemudian dilakukan perhitungan berdasarkan metode *K-NN* dengan *Euclidean distance* sebagai berikut :

No. : 11234

IPS 1 : 3,65

IPS 2 : 3,67

IPS 3 : 2,83

Dari data baru untuk mahasiswa dengan nomor 11234 tersebut dihitung jarak dengan menggunakan rumus (2.2) nilai $K=5$ sebagai berikut:

1. Nilai $d = (x_1 - y_1)^2$

$$\begin{aligned} & \sqrt{(3,65 - 3,48)^2 + (3,67 - 3,65)^2 + (2,83 - 3,38)^2} \\ &= \sqrt{0,0289 + 0,0004 + 0,3025} \\ &= \sqrt{0,3318} \\ &= 0,576020833 \end{aligned}$$

2. Nilai $d = (x_1 - y_2)^2$

$$\begin{aligned} & \sqrt{(3,65 - 3,13)^2 + (3,67 - 3,14)^2 + (2,83 - 3,00)^2} \\ &= \sqrt{0,2704 + 0,2809 + 0,0289} \\ &= \sqrt{0,5802} \\ &= 0,7617086057 \end{aligned}$$

3. Nilai $d = (x_1 - y_3)^2$

$$\begin{aligned} & \sqrt{(3,65 - 3,13)^2 + (3,67 - 3,14)^2 + (2,83 - 3,00)^2} \\ &= \sqrt{0,1681 + 0,3481 + 0} \\ &= \sqrt{0,5162} \\ &= 0,7184705979 \end{aligned}$$

Berikut ini adalah table hasil perhitungan jarak menggunakan *Euclidean Distance* :

Tabel 4. 3 Perhitungan Jarak Menggunakan *Euclidean Distance*

No.	Jarak	Selisih
2291	0,576020833	5 Tahun
10311	0,642028368	4,5 Tahun
3201	0,7184705979	5 Tahun
4531	0,7491995729	4 Tahun
2951	0,7617086057	5 Tahun
8421	0,76223356	4,5 Tahun
4532	0,9811727677	4 Tahun
8031	1,0507140429	5,5 Tahun
10381	1,1792370415	4 Tahun
8540	1,1942361576	4,5 Tahun

Dari data yang telah diurutkan di atas menggunakan $K=5$ untuk mengambil data dengan nilai jarak terkecil.

Tabel 4. 4 Jarak Terdekat sebanyak $K=5$

No.	Jarak	Selisih
2291	0,576020833	5 Tahun
10311	0,642028368	4,5 Tahun
3201	0,7184705979	5 Tahun
4531	0,7491995729	4 Tahun
2951	0,7617086057	5 Tahun

Lakukan evaluasi terhadap selisih data dan tentukan selisih lama masa studi yang memiliki jumlah terbanyak, maka selisih dengan jumlah terbanyak tersebut adalah hasil klasifikasi, dan didapatkan 5 tahun sebanyak 3 data, 4,5 tahun 1 data dan 4 tahun 1 data. Maka dapat disimpulkan lama masa studi yang akan ditempuh oleh mahasiswa bernomor 11234 yaitu 5 tahun.

4.2.2 Evaluasi Model Dengan *Mean Square Error (MSE)*

Setelah dilakukan pemodelan algoritma *K-Nearest Neighbor*, selanjutnya dilakukan pengujian akurasi dengan mencari nilai *error* antara data aktual dan data prediksi menggunakan *Mean Square Error (MSE)*.

Nilai MSE didapat dari persamaan :

$$\text{MSE} = \frac{\sum_{t=1}^n (y_t - y'_t)^2}{n}$$

Keterangan:

y_t = data aktual pada periode t

y'_t = data peramalan pada periode t

n = jumlah data

1. $2291 = (y_t - y'_t)^2$
 $= (5 - 5)^2 = 0$
2. $2951 = (y_t - y'_t)^2$
 $= (5 - 5)^2 = 0$
3. $3201 = (y_t - y'_t)^2$
 $= (5 - 5)^2 = 0$
4. $4531 = (y_t - y'_t)^2$
 $= (4 - 5)^2 = 1$
5. $4532 = (y_t - y'_t)^2$
 $= (4 - 5)^2 = 1$
6. $8031 = (y_t - y'_t)^2$
 $= (5 - 5)^2 = 0,25$
7. $8421 = (y_t - y'_t)^2$
 $= (5,5 - 5)^2 = 0,25$
8. $8540 = (y_t - y'_t)^2$
 $= (4,5 - 5)^2 = 0,25$
9. $10311 = (y_t - y'_t)^2$
 $= (4,5 - 5)^2 = 0,25$
10. $10381 = (y_t - y'_t)^2$
 $= (5 - 5)^2 = 1$

Kemudian jumlahkan hasil nilai errornya dan dibagi dengan jumlah data yang ada :

$$0 + 0 + 0 + 1 + 1 + 0,25 + 0,25 + 0,25 + 0,25 + 1 = 4$$

$$\frac{4}{10} = 0,4$$

Tabel 4. 5 Evaluasi *Mean Square Error*

No.	Lama Studi (y_t)	Jarak	Rangking	K=5	Hasil Peramalan (y'_t)	<i>Mean Square Error</i>
2291	5	0,576020833	1	5 Tahun	5	0
2951	5	0,761708606	5	5 Tahun	5	0
3201	5	0,718470598	3	5 Tahun	5	0
4531	4	0,749199573	4	4 Tahun	5	1
4532	4	0,974217635	7		5	1
8031	5,5	1,060235823	8		5	0,25
8421	4,5	0,76223356	6		5	0,25
8540	4,5	1,194236158	10		5	0,25
10311	4,5	0,642028037	2	4,5 Tahun	5	0,25
10381	4	1,179237041	9		5	1
				5 Tahun		0,4

Dari hasil evaluasi model menggunakan *Mean Square Error (MSE)*, penerapan algoritma *K-Nearest Neighbor (K-NN)* pada prediksi lama masa studi mahasiswa mendapat hasil 0,4 sebagai *errornya*.

BAB V

PEMBAHASAN HASIL

5.1 Objek dan Lokasi Penelitian

5.1.1 Profil Universitas Ichsan Gorontalo

Universitas Ichsan Gorontalo dengan SK No 84/D/0/2001 dengan Alamat Kampus 1 Di Jalan Achmad Nadjamuddin No. 17, dan Kampus 2 Di Jalan Trans Sulawesi No, 147 Provinsi Gorontalo dengan No Telepon (0435) 829975 / (0435) 829976. Memiliki website dengan alamat <http://www.unisan.ac.id> dengan motto kampus yaitu “ *quality is our target* “ .

Universitas Ichsan Gorontalo atau lebih dikenal dengan kampus hijau merupakan perguruan tinggi di Indonesia Timur yang unggul dan terdepan dalam penyelenggaraan tridarma perguruan tinggi untuk menghasilkan lulusan berkualitas, inovatif, kreatif, intrepeneur yang memiliki semangat kerakyatan, pengabdian dan kemandirian dalam mengembangkan ilmu dan teknologi yang mampu mendukung pembangunan nasional, memajukan dirinya dan masyarakat, mempertinggi mutu kebudayaan sebagai bagian dari kehidupan masyarakat.

Universitas Ichsan Gorontalo ialah satu dari sekian kampus swasta di Indonesia yang berupa Universitas, diurus oleh dikti dan tercatat kedalam kopertis wilayah 9. Kampus ini telah berdiri sejak tahun 10 juli 2001 dengan nomor PT 84do2001 dan tanggal SK PT 10 Juli 2001.

Universitas Ichsan Gorontalo bermula dari sebuah lembaga kursus komputer, kemudian berkembang menjadi program setara Diploma 1 (D1) dan selanjutnya menjadi sekolah tinggi dan kini menjadi salah satu universitas terbesar di Gorontalo yang membina program Strata 1 atau (S1) dan mulai ini sudah dibuka program pasca sarjana (S2).

Universitas Ichsan Gorontalo merupakan universitas pertama dikawasan Indonesia bagian timur yang memberikan mata kuliah komputer sebagai mata kuliah wajib bagi seluruh mahasiswa pada jurusan atau fakultas yang ada.

Di samping itu, Universitas Ichsan Gorontalo juga membina suasana kehidupan akademik yang sehat dan bertanggung jawab serta mengembangkan dan mengusahakan temuan baru dalam ilmu pengetahuan dan teknologi melalui penggunaan sumber daya yang ada secara maksimal untuk mencapai *Good Ichsan University Governance*.

5.1.2 Sejarah Singkat Universitas Ichsan Gorontalo

Sejak tahun 1986, untuk pertama kalinya Dr. Abd Gaffar Latjoke, M.Si Pertama menginjakkan kaki di Gorontalo selaku dosen di IKIP Gorontalo, Kini beralih menjadi Universitas Negeri Gorontalo perjalanan waktu kemudian menempatkan dirinya kini menjadi Rektor Universitas Ichsan Gorontalo salah satu Universitas swasta yang menjadi pilihan cukup banyak generasi muda Gorontalo dan sekitarnya dalam melanjutkan studi jenjang perguruan tinggi.

Universitas Ichsan Gorontalo pada awalnya yakni tahun 1987 hanya sebuah lembaga Kursus akuntansi. mengetik dan bahasa inggris membuka lembaga kursus karena memenuhi tuntutan dari masyarakat akan tenaga terampil untuk kantor pemerintah dan swasta.

Pria kelahiran Lawawoi Sidrap 31 Desember 1962 itu kemudian mengambil peluang yang ada dengan membuka lembaga kursus yang diberi nama Ichsan yang juga merupakan anak pertamanya. Dengan tuntutan masyarakat yang semakin kuat, lembaga kursus tersebut kemudian di ubah menjadi program Diploma 1 (D1) Komputer tahun 2000 kembali diubah menjadi Stimik dan Stie Icshan.

Lembaga pendidikan itu berubah menjadi Universitas Ichsan Gorontalo pada tahun 2001 dengan membuka 6 fakultas yakni fakultas hukum, Ekonomi, Ilmu Sosial dan Ilmu Politik, Teknik dan Ilmu Komputer.

5.2 Pembahasan Hasil Eksperimen

Pembahasan hasil eksperimen adalah sebuah langkah terakhir yang ditulis peneliti untuk menjelaskan hasil pembahasan pada bab sebelumnya. Langkah ini diambil karena menurut peneliti sesuai dengan hasil eksperimen harus dijelaskan dalam pembahasan hasil eksperimen. Pembahasan hasil eksperimen meliputi

pengumpulan data, pra pengolahan data, penerapan metode dan evaluasi model. Oleh karena itu pembahasan hasil eksperimen ini menjadi hal yang penting dituliskan untuk menunjukkan seberapa besar hasil yang di bab sebelumnya.

5.2.1 Pembahasan Pengumpulan Data

Dalam pembahasan pengumpulan data yang didapatkan oleh penelitian ini diambil atau diperoleh dari BAAK Universitas Ichsan Gorontalo sampai dengan Juni 2021. Data alumni yang diperoleh adalah sebanyak 10397 alumni mahasiswa Universitas Ichsan Gorontalo. Sesuai dengan jumlah data yang diperoleh, lama masa studi terdapat berbagai lama masa studi. Diantaranya 4 tahun, 4,5 tahun, 5 tahun, 5,5 tahun, dan ada juga sampai dengan 7 tahun

5.2.2 Pembahasan Pra Pengolahan Data

Dalam pembahasan pra pengolahan data pada bab sebelumnya peneliti menemukan adanya *missing value* dan *outlier*, yang akan dijelaskan sebagai berikut :

1. Missing Value

Missing value dalam ilmu statistik adalah data yang hilang yang terjadi ketika tidak ada nilai data yang disimpan untuk variabel dalam pengamatan. Data yang hilang adalah kejadian umum dan dapat memiliki efek signifikan pada kesimpulan yang dapat dari data.

Dari data yang ada, terdapat beberapa *missing value* yang telah dijelaskan pada bab sebelumnya, yaitu :

- a. Indeks Prestasi Semester 1 : 46
- b. Indeks Prestasi Semester 2 : 160
- c. Indeks Prestasi Semester 3 : 312
- d. Masa Studi : 991

2. *Outlier* (nilai tidak wajar)

Outlier merupakan suatu observasi pada kumpulan data yang berbeda nilainya dari observasi-observasi lainnya pada kumpulan data tersebut. Sesuai dengan penarikan data menggunakan *Interquartile Range (IQR)* dapat dijelaskan sebagai berikut :

- a. Pada **Gambar 4.1** dan **Gambar 4.2** menampilkan Indeks Prestasi Semester 1 dan Indeks Prestasi Semester 2 terdapat nilai yang tidak wajar yaitu nilai 0 sampai dengan nilai 2.
- b. Pada **Gambar 4.3** terdapat lama masa studi yang tidak wajar yaitu lebih dari 5,5 tahun.

5.2.3 Hasil Penerapan Algoritma *K-Nearest Neighbor (K-NN)*

Algoritma *K-Nearest Neighbor (KNN)* bekerja berdasarkan jarak terpendek dari *query instance* ke training data untuk menentukan *KNN*-nya. Salah satu cara untuk menghitung jarak dekat atau jauhnya tetangga menggunakan metode *euclidean distance*.

Setelah diterapkan Algoritma *K-Nearest Neighbor* maka untuk memprediksi lama masa studi mahasiswa melalui Langkah mencantumkan indeks prestasi semester 1, 2 dan 3. Berikut juga adalah hasil evaluasi terhadap selisih data dan tentukan selisih lama masa studi yang memiliki jumlah terbanyak, maka selisih dengan jumlah terbanyak tersebut adalah hasil klasifikasi, dan didapatkan 5 tahun sebanyak 3 data, 4,5 tahun 1 data dan 4 tahun 1 data. Maka dapat disimpulkan lama masa studi yang akan ditempuh oleh mahasiswa bernomor 11234 yaitu 5 tahun

5.2.4 Pembahasan Hasil *Mean Square Error (MSE)*

Mean square Error adalah rata-rata dari kesalahan *forecast* dikuadratkan. *Mean Square Error (MSE)* menggunakan nilai kuadrat untuk setiap selisih perhitungan yang terjadi. Pada **Tabel 4.5** adalah model pencarian nilai *error* yang didapatkan setelah peneliti meramalkan lama masa studi mahasiswa menggunakan algoritma *K-Nearest Neighbor (K-NN)*.

BAB VI

KESIMPULAN DAN SARAN

6.1 Kesimpulan

Berdasarkan hasil penelitian ini, peneliti menyimpulkan bahwa :

Penerapan metode *K-Nearest Neighbor (K-NN)* dalam memprediksi lama masa studi mahasiswa di Universitas Ichsan Gorontalo berhasil dilakukan. Dengan gambaran sederhana hanya dengan memasukkan indeks prestasi semester 1, indeks prestasi semester 2 dan indeks prestasi semester 3, lama masa studi mahasiswa sudah bisa diprediksi. Penggunaan metode *K-Nearest Neighbor (K-NN)* layak digunakan dalam lingkup Universitas Ichsan Gorontalo guna memprediksi lama masa studi mahasiswa.

6.2 Saran

Adapun saran berdasarkan penelitian ini berupa :

1. Untuk menangani *missing value* pada data yang ada, dibutuhkan metode yang tepat.
2. Dibutuhkan metode *interquartile range (IQR)* untuk menarik kesimpulan mengenai sekumpulan data *outlier*.

DAFTAR PUSTAKA

- [1] I. dan A. Mutiara, “Penerapan K-Optimal Pada Algoritma Knn Untuk Prediksi Kelulusan Tepat Waktu Mahasiswa Program Studi Ilmu Komputer Fmipa Unlam Berdasarkan Ip Sampai Dengan Semester 4,” *Klik - Kumpul. J. Ilmu Komput.*, vol. 2, no. 2, pp. 159–173, 2015.
- [2] L. Data *et al.*, “Algoritma K-Nearest Neighbor Classification Sebagai Sistem Prediksi Predikat Prestasi Mahasiswa,” vol. 13, no. 2, pp. 195–202, 2016.
- [3] E. S. Susanto and H. Al Fatta, “Informatika Universitas Amikom Yogyakarta Menggunakan Metode K-Nearest Neighbor,” vol. Xiii, pp. 67–72, 2018.
- [4] F. Rohmawati, M. G. Rohman, and S. Mujilahwati, “Sistem Prediksi Jumlah Pengunjung Wisata Wego Kec.Sugio Kab.Lamongan Menggunakan Metode Fuzzy Time Series,” *Jouticla*, vol. 2, no. 2, 2017, doi: 10.30736/jti.v2i2.66.
- [5] A. D. Sinaga, “Prediksi Kelulusan Mahasiswa Fakultas Sains Dan Teknologi Universitas Sanata Dharma Menggunakan Metode Klasifikasi Naïve Bayes,” p. 66, 2020, [Online]. Available: repository.usd.ac.id.
- [6] R. M. Abarca, “Peraturan Akademik Universitas Ichsan Gorontalo,” *Nuevos Sist. Comun. e Inf.*, pp. 2013–2015, 2019.
- [7] A. K. Neighborhood, “Pembuatan Aplikasi Data Mining Untuk Pendahuluan Latar Belakang Maksud Dan Tujuan Batasan Masalah,” pp. 1–29, 2012.
- [8] D. Mining and D. Mining, “Bab 2 Tinjauan Pustaka Dan Dasar Teori 2.1,” no. 2015, pp. 7–17, 2014.
- [9] A. E. Armi, A. H. Kridalaksana, and Z. Arifin, “Peramalan Angka Inflasi Kota Samarinda Menggunakan Metode Double Exponential Smoothing (Studi Kasus : Badan Pusat Statistik Kota Samarinda),” *Inform. Mulawarman J. Ilm. Ilmu Komput.*, vol. 14, no. 1, p. 21, Feb. 2019, doi: 10.30872/jim.v14i1.1252.

KODE PROGRAM

```
import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

import seaborn as sns

import missingno as mno

from sklearn import linear_model


df = pd.read_csv("grad_dataset.csv")

df.head()

df.describe(include="all")

df.dtypes

df['Masuk'] = df['Masuk'].replace('\n',' ', regex=True)

df['Lulus'] = df['Lulus'].replace('\n',' ', regex=True)

df.head()

df['Masuk'] = df['Masuk'].replace('Januari','January', regex=True)

df['Masuk'] = df['Masuk'].replace('Februari','February', regex=True)

df['Masuk'] = df['Masuk'].replace('Maret','March', regex=True)

df['Masuk'] = df['Masuk'].replace('April','April', regex=True)

df['Masuk'] = df['Masuk'].replace('Mei','May', regex=True)

df['Masuk'] = df['Masuk'].replace('Juni','June', regex=True)

df['Masuk'] = df['Masuk'].replace('Juli','July', regex=True)
```

```
df['Masuk'] = df['Masuk'].replace('Agustus','August', regex=True)
df['Masuk'] = df['Masuk'].replace('September','September', regex=True)
df['Masuk'] = df['Masuk'].replace('Oktober','October', regex=True)
df['Masuk'] = df['Masuk'].replace('Nopember','November', regex=True)
df['Masuk'] = df['Masuk'].replace('Desember','December', regex=True)
df['Masuk'] = df['Masuk'].replace('00 January 1900','NaN', regex=True)
```

```
df['Lulus'] = df['Lulus'].replace('Januari','January', regex=True)
df['Lulus'] = df['Lulus'].replace('Februari','February', regex=True)
df['Lulus'] = df['Lulus'].replace('Maret','March', regex=True)
df['Lulus'] = df['Lulus'].replace('April','April', regex=True)
df['Lulus'] = df['Lulus'].replace('Mei','May', regex=True)
df['Lulus'] = df['Lulus'].replace('Juni','June', regex=True)
df['Lulus'] = df['Lulus'].replace('Juli','July', regex=True)
df['Lulus'] = df['Lulus'].replace('Agustus','August', regex=True)
df['Lulus'] = df['Lulus'].replace('September','September', regex=True)
df['Lulus'] = df['Lulus'].replace('Oktober','October', regex=True)
df['Lulus'] = df['Lulus'].replace('Nopember','November', regex=True)
df['Lulus'] = df['Lulus'].replace('Desember','December', regex=True)
df['Lulus'] = df['Lulus'].replace('00 January 1900','NaN', regex=True)
```

```
df.head()
```

```
df.isna().sum()

df['Masuk'] = pd.to_datetime(df['Masuk'])

df['Lulus'] = pd.to_datetime(df['Lulus'])

df['Tanggal Lahir'] = pd.to_datetime(df['Tanggal Lahir'], format='%m/%d/%Y')

df.head()

df.dtypes

df.describe(include='all')

df['Usia'] = round((df['Masuk'] - df['Tanggal Lahir']).astype('<m8[M]')/12)

df['Masa Studi'] = round((df['Lulus'] - df['Masuk']).astype('<m8[M]')/12,2)

df.head()

pd_df = df.drop(['Masuk','Lulus','Tanggal Lahir'],axis=1)

pd_df.head()

pd_df.describe()

pd_df.notna()

pd_df.isna().sum()

mno.matrix(pd_df, figsize = (20, 6))

modus_value = pd_df['Kelamin'].mode()[0]

pd_df['Kelamin'] = pd_df['Kelamin'].fillna(modus_value)

from sklearn.preprocessing import LabelEncoder

ProdiEnc = LabelEncoder()

ProdiEnc.fit(pd_df['Prodi'])

pd_df['Prodi'] = ProdiEnc.transform(pd_df['Prodi'])
```



```

ShiftEnc = LabelEncoder()

ShiftEnc.fit(pd_df['Shift'])

pd_df['Shift'] = ShiftEnc.transform(pd_df['Shift'])


genderEnc =LabelEncoder()

genderEnc.fit(pd_df['Kelamin'])

pd_df['Kelamin'] = genderEnc.transform(pd_df['Kelamin'])

missing_columns = ["IPS_1", "IPS_2","IPS_3","Usia", "Masa Studi"]

def random_imputation(pd_df, feature):

    number_missing = pd_df[feature].isnull().sum()

    observed_values = pd_df.loc[pd_df[feature].notnull(), feature]

    pd_df.loc[pd_df[feature].isnull(), feature + '_imp'] =
np.random.choice(observed_values, number_missing, replace = True)


    return pd_df

for feature in missing_columns:

    pd_df[feature + '_imp'] = pd_df[feature]

    pd_df = random_imputation(pd_df, feature)

deter_data = pd.DataFrame(columns = ["Det" + name for name in
missing_columns])

for feature in missing_columns:

```

```

deter_data["Det" + feature] = pd_df[feature + "_imp"]

parameters = list(set(pd_df.columns) - set(missing_columns) - {feature +
'_imp'})

model = linear_model.LinearRegression()

model.fit(pd_df[parameters], pd_df[feature + '_imp'])

deter_data.loc[pd_df[feature].isnull(), "Det" + feature] =
model.predict(pd_df[parameters])[pd_df[feature].isnull()]

df_nomissing_value = pd.concat([deter_data, pd_df[['Prodi', 'Shift', 'Kelamin']]],
axis=1)

df_nomissing_value.rename(columns={"DetMasa Studi": "Masa Studi",
"DetUsia": "Usia", "DetIPS_1": "IPS_1",
"DetIPS_2": "IPS_2", "DetIPS_3": "IPS_3"},

inplace=True)

df_nomissing_value

newdf_final = df_nomissing_value[df_nomissing_value['Masa Studi']>3.8]

newdf_final.describe()

import warnings

warnings.filterwarnings("ignore")

sns.set()

fig, axes = plt.subplots(nrows=3, ncols=2)

fig.set_size_inches(10, 10)

for index, variable in enumerate(['IPS_1', 'IPS_2', 'Masa Studi']):

sns.boxplot(newdf_final[variable], x=variable, ax = axes[index, 0])

sns.distplot(newdf_final[variable], ax = axes[index, 1])

```

```

plt.tight_layout()

q1 = (newdf_final[['IPS_1','IPS_2', 'Masa Studi']]).quantile(0.25)

q3 = (newdf_final[['IPS_1','IPS_2', 'Masa Studi']]).quantile(0.75)

IQR = q3 - q1

maximum = q3 + (1.5 * IQR)

minimum = q1 - (1.5 * IQR)


print(maximum)

print(minimum)

batas_bawah_IPS_1 = (newdf_final['IPS_1'] < minimum[0])

newdf_final['IPS_1'] = newdf_final['IPS_1'].mask(batas_bawah_IPS_1,
minimum[0])

batas_atas_IPS_1 = (newdf_final['IPS_1'] > 4.00)

newdf_final['IPS_1'] = newdf_final['IPS_1'].mask(batas_atas_IPS_1, 4.00)

batas_bawah_IPS_2 = (newdf_final['IPS_2'] < minimum[1])

newdf_final['IPS_2'] = newdf_final['IPS_2'].mask(batas_bawah_IPS_2,
minimum[1])

batas_atas_IPS_2 = (newdf_final['IPS_2'] > 4.00)

newdf_final['IPS_2'] = newdf_final['IPS_2'].mask(batas_atas_IPS_2, 4.00)

#batas_bawah_masa_studi = (newdf_final['Masa Studi'] < minimum[2])

#newdf_final['Masa Studi'] = newdf_final['Masa
Studi'].mask(batas_bawah_masa_studi, minimum[2])

batas_atas_masa_studi = (newdf_final['Masa Studi'] > maximum[2])

```

```

newdf_final['Masa Studi'] = newdf_final['Masa
Studi'].mask(batas_atas_masa_studi, maximum[2])

newdf_final.describe()

import warnings

warnings.filterwarnings("ignore")

sns.set()

fig, axes = plt.subplots(nrows=3, ncols=2)

fig.set_size_inches(10, 10)

for index, variable in enumerate(['IPS_1', 'IPS_2', 'Masa Studi']):

    sns.boxplot(newdf_final[variable], x=variable, ax = axes[index, 0])

    sns.distplot(newdf_final[variable], ax = axes[index, 1])

plt.tight_layout()

SpearmanCorr = newdf_final.corr(method="spearman")

sns.heatmap(SpearmanCorr, vmax=.9, square=True, annot=True, linewidths=.3,
cmap="YlGnBu", fmt='.1f')

df_finalfeature = newdf_final[["IPS_1", "IPS_2", "IPS_3"]]

df_finaltarget = newdf_final[['Masa Studi']]

from sklearn.preprocessing import StandardScaler, MinMaxScaler

sc_feature = StandardScaler()

X = sc_feature.fit_transform(df_finalfeature.values)

#sc_target = StandardScaler()

#y = sc_target.fit_transform(df_finaltarget.values)

```

```

import pickle

with open('feature_model.pkl', 'wb') as files:

    pickle.dump(sc_feature, files)

from sklearn.model_selection import train_test_split

x_train, x_test, y_train, y_test = train_test_split(X, df_finaltarget.values,
test_size=0.3, random_state=4)

from sklearn import neighbors

from sklearn.model_selection import RandomizedSearchCV

grid = dict(n_neighbors = [2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20])

searcher = RandomizedSearchCV(neighbors.KNeighborsRegressor(),

                             param_distributions = grid, cv=10)

search_result = searcher.fit(x_train, y_train)

bestScore = search_result.best_score_

bestParams = search_result.best_params_

bestModel = search_result.best_estimator_

y_pred = bestModel.predict(x_test)


from sklearn.metrics import r2_score

score_MAE_KNN = round(np.mean(np.absolute(y_pred - y_test)),2)

print("KNN for Regression MAE: %.2f" % score_MAE_KNN)

score_MSE_KNN = round(np.mean((y_pred - y_test)** 2),2)

```

```
print("KNN for Regression MSE: %.2f" % score_MSE_KNN)

result_KNN = ['KNN for Regression', score_MAE_KNN,
score_MSE_KNN]filename = "resultmodel.sav"

pickle.dump(bestModel, open(filename, 'wb'))
```




**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 : 034/Perpustakaan-Fikom/VI/2022

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

Nama Anggota : Faruq Natsir Djafar

No. Induk : T3117023

No. Anggota : M202275

Terhitung mulai hari, tanggal : Jumat, 10 Juni 2022, dinyatakan telah bebas pinjam buku dan koleksi perpustakaan lainnya.

Demikian keterangan ini di buat untuk di gunakan sebagaimana mestinya.

Gorontalo, 10 Juni 2022

**Mengetahui,
Kepala Perpustakaan**



Apriyanto Alhamad, M.Kom

NIDN : 0924048601



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN TEKNOLOGI
UNIVERSITAS ICHSAN GORONTALO
LEMBAGA PENELITIAN

Kampus Unisan Gorontalo Lt.3 - Jln. Achmad Nadjamuddin No. 17 Kota Gorontalo
Telp. (0435) 8724466, 829975 E-Mail: lembagapenelitian@unisan.ac.id

Nomor : 3884/PIP/LEMLIT-UNISAN/GTO/II/2022

Lampiran : -

Hal : Permohonan Izin Penelitian

Kepada Yth,

Kepala BAAK Universitas Ichsan Gorontalo

di,-

Tempat

Yang bertanda tangan di bawah ini :

Nama : Dr. Rahmisyari, ST.,SE.,MM

NIDN : 0929117202

Jabatan : Ketua Lembaga Penelitian

Meminta kesediannya untuk memberikan izin pengambilan data dalam rangka penyusunan **Proposal / Skripsi**, kepada :

Nama Mahasiswa : Faruq Natsir Djafar

NIM : T3117023

Fakultas : Fakultas Ilmu Komputer

Program Studi : Teknik Informatika

Lokasi Penelitian : BAAK UNIVERSITAS ICHSAN GORONTALO

Judul Penelitian : PENERAPAN ALGORITMA K-NEAREST NEIGHBOR (K-NN) UNTUK PREDIKSI LAMA MASA STUDI MAHASISWA DI UNIVERSITAS ICHSAN GORONTALO

Atas kebijakan dan kerja samanya diucapkan banyak terima kasih.

Gorontalo, 15 Februari 2022
Ketua

Dr. Rahmisyari, ST.,SE.,MM
NIDN 0929117202



**KEMENTERIAN PENDIDIKAN, KEBUDAYAAN,
RISET DAN TEKNOLOGI
UNIVERSITAS ICHSAN
(UNISAN) GORONTALO**

SURAT KEPUTUSAN MENDIKNAS RI NO. 84/D/O/2001
JL. Achmad Nadjamuddin No. 17 Telp. (0435) 829975 Fax (0435) 829976 Gorontalo

SURAT KETERANGAN
NO : 090/UNISAN-G/VI/2022

Yang bertanda tangan dibawah ini :

Nama : Amiruddin, M.Kom
NIDN : 0910097601
Jabatan : Wakil Rektor Bidang Akademik, Tata Kelola & Sistem Informasi

Dengan ini menerangkan bahwa :

Nama : Faruq Natsir Djafar
Nim : T3117023
Prodi : Teknik Informatika
Fakultas : Ilmu Komputer
Universitas : Ichsan Gorontalo

Adalah benar telah melakukan penelitian dalam rangka penulisan Skripsinya yang berjudul
"Penerapan Algoritma K-Nearest Neighbor (K-NN) Untuk Prediksi Lama Masa Studi
Mahasiswa Di Universitas Ichsan Gorontalo" sejak tanggal 26 November – 12 Februari 2022,
dan telah pula membahas materi hasil penelitiannya dengan kami.

Demikian surat keterangan ini dibuat untuk dapat dipergunakan sebagaimana mestinya,

Gorontalo, 14 Juni 2022

An. Rektor
Wakil Rektor Bidang Akademik, Tata Kelola &
Sistem Informasi

Amiruddin, M.Kom
NIDN. 0910097601

Tembusan:

1. Rektor Universitas Ichsan Gorontalo
2. Ketua Lemlit Universitas Ichsan Gorontalo
3. Mahasiswa yang bersangkutan
4. Arsip

DAFTAR RIWAYAT HIDUP

A. IDENTITAS PRIBADI



Nama	: Faruq Natsir Djafar
Nim	: T3117023
Tempat, Tanggal Lahir	: Kotamobagu, 23 April 1999
Agama	: Islam
Email	: faruqnatsir234@gmail.com

B. RIWAYAT PENDIDIKAN

1. Tahun 2011, Menyelesaikan Pendidikan di Sekolah Dasar Negeri 2 Gogagoman, Kec. Kotamobagu Barat, Kota Kotamobagu.
2. Tahun 2014, Menyelesaikan Pendidikan di Madrasah Tsanawiyah Negeri Kotamobagu, Kota Kotamobagu.
3. Tahun 2017, Menyelesaikan Pendidikan di Madrasah Aliyah Negeri 1 Kotamobagu, Kota Kotamobagu.
4. Tahun 2017, Telah Diterima Menjadi Mahasiswa Di Perguruan Tinggi Swasta Universitas Ichsan Gorontalo.

PAPER NAME

SKRIPSI_T3117023_FARUQ NATSIR DJA
FAR.docx

AUTHOR

Faruq Natsir Djafar faruqnatsir234@gmail.com

WORD COUNT

5102 Words

CHARACTER COUNT

30992 Characters

PAGE COUNT

35 Pages

FILE SIZE

237.0KB

SUBMISSION DATE

Jun 13, 2022 1:03 PM GMT+8

REPORT DATE

Jun 13, 2022 1:05 PM GMT+8

● **29% Overall Similarity**

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

- 29% Internet database
- 5% Publications database
- Crossref database
- Crossref Posted Content database
- 2% Submitted Works database

● **Excluded from Similarity Report**

- Bibliographic material
- Small Matches (Less than 25 words)

29% Overall Similarity

Top sources found in the following databases:

- 29% Internet database
- 5% Publications database
- Crossref database
- Crossref Posted Content database
- 2% Submitted Works database

TOP SOURCES

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1	eprints.akakom.ac.id Internet	7%
2	repository.usd.ac.id Internet	4%
3	text-id.123dok.com Internet	3%
4	cosphijournal.unisan.ac.id Internet	2%
5	spada.unisan.ac.id Internet	2%
6	yahyaunsa.blogspot.com Internet	2%
7	docobook.com Internet	1%
8	123dok.com Internet	1%

9	jti.respati.ac.id	Internet	1%
10	jurnal.unprimdn.ac.id	Internet	1%
11	journal.upgris.ac.id	Internet	<1%
12	kaskus.co.id	Internet	<1%
13	akupintar.id	Internet	<1%
14	ejournal.itn.ac.id	Internet	<1%
15	scribd.com	Internet	<1%
16	conference.binadarma.ac.id	Internet	<1%
17	repository.unmuhjember.ac.id	Internet	<1%