

**IMPLIMENTASI DATA MINING PADA PENILAIAN KINERJA GURU
MENGUNAKAN METODE K-NEAREST NEIGHBOR (K-NN) STUDI
KASUS PADA MI ALFAGIRI SILO-JEMBER**

SKRIPSI



Oleh:

MUHAMMAD JUNDANUDDIN

NPM : 2020502025

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS SAINS DAN TEKNOLOGI UNIVERSITAS IBRAHIMY
SITUBONDO**

2024

**IMPLIMENTASI DATA MINING PADA PENILAIAN KINERJA GURU
MENGUNAKAN METODE K-NEAREST NEIGHBOR (K-NN) STUDI
KASUS PADA MI ALFAGIRI SILO-JEMBER**

SKRIPSI



Oleh:

MUHAMMAD JUNDANUDDIN

NPM : 2020502025

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS SAINS DAN TEKNOLOGI UNIVERSITAS IBRAHIMY
SITUBONDO**

2024

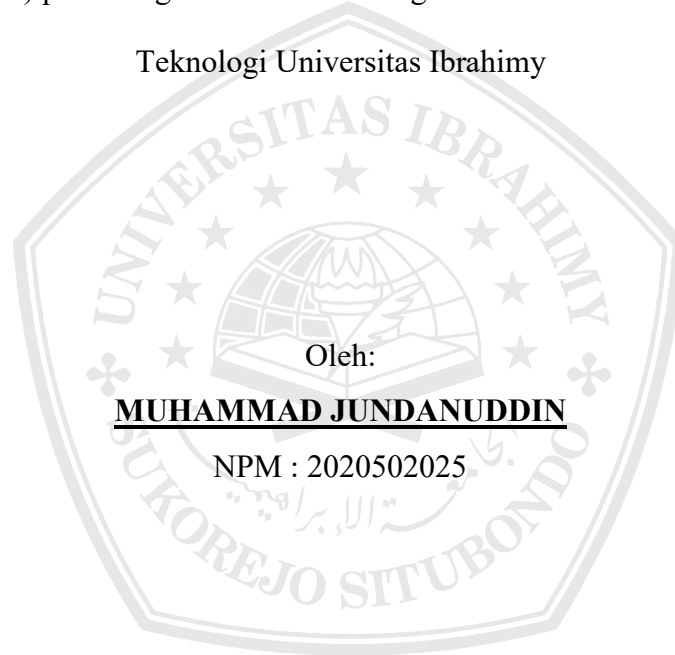
**IMPLIMENTASI DATA MINING PADA PENILAIAN KINERJA
GURU MENGGUNAKAN METODE K-NEAREST NEIGHBOR (K-
NN) STUDI KASUS PADA MI ALFAGIRI SILO-JEMBER**

SKRIPSI

Diajukan untuk Memenuhi Salah Satu Persyaratan dalam Menyelesaikan Program

Sarjana (S-1) pada Program Studi Teknologi Informasi Fakultas Sains dan

Teknologi Universitas Ibrahimy



Oleh:

MUHAMMAD JUNDANUDDIN

NPM : 2020502025

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS SAINS DAN TEKNOLOGI UNIVERSITAS IBRAHIMY
SITUBONDO**

2024

PERNYATAAN KEASLIAN TULISAN

Saya yang bertanda tangan di bawah ini:

Nama : **Muhammad Jundanuddin**

NPM/NIRM : 2020502025

Program Studi : S-1 Sistem Informasi

Fakultas : Fakultas Sains dan Teknologi Universitas Ibrahimi

Menyatakan dengan sebenarnya, bahwa tugas akhir/skripsi ini secara keseluruhan adalah hasil penelitian atau karya saya sendiri, kecuali pada bagian-bagian yang dirujuk sebagai sumber referensi dan disebutkan dalam daftar pustaka. Apabila di kemudian hari terbukti atau dapat dibuktikan bahwa tugas akhir/skripsi ini hasil plagiasi, maka saya bersedia menerima sanksi atas perbuatan tersebut.



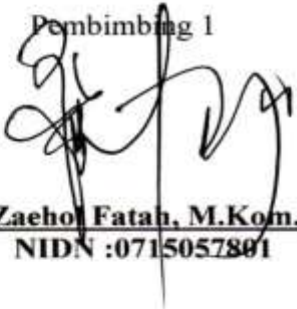
Muhammad Jundanuddin
NPM : 2020502025

PERSETUJUAN PEMBIMBING

Nama : **Muhammad Jundanuddin**
NPM/NIRM : 2020502025
Judul : **IMPLIMENTASI DATA MINING PADA PENILAIAN
KINERJA GURU MENGGUNAKAN METODE K-
NEAREST NEIGHBOR (K-NN) STUDI KASUS PADA
MI ALFAGIRI SILO-JEMBER**

Telah disetujui oleh :

Pembimbing 1


Zaehol Fatah, M.Kom.
NIDN : 0715057891

Pembimbing 2


Achmad Baijuri, M.Kom.
NIDN : 0715078902

PERSETUJUAN PEMBIMBING

PENGESAHAN

SKRIPSI

**IMPLIMENTASI DATA MINING PADA PENILAIAN KINERJA GURU
MENGUNAKAN METODE K-NEAREST NEIGHBOR (K-NN) STUDI
KASUS PADA MI ALFAGIRI SILO-JEMBER**


MUHAMMAD JUNDANUDDIN

2020502025


telah dipertahankan didepan dewan penguji Sidang/Munaqosyah Skripsi pada hari
Kamis Tanggal 15 Agustus 2024 sebagai salah satu syarat memperoleh gelar
Sarjana (S.Kom) pada Fakultas Sains dan Teknologi Universitas Ibrahimy.

Tim Penguji,


Ketua Sidang,


Abdul Wafi, S. Pi, M.P.
NIDN.0705049103

Sekretaris Sidang,


Abdus Samad, M.Kom.
NIDN. 0709099006

Penguji I,


Lukman Fakhri Lidinillah, M.Kom.
NIDN.0715099001

Penguji II,


Akhlis Munazilin, S.Kom., M.T
NIDN.0712098601

Mengetahui
Dekan,


Abd. Ghofur, M.Kpm.
NIDN.0711088303

PERSEMBAHAN

Saya persembahkan karya kecil ini kepada kedua orang tua dan keluarga ku yang selalu membuatku termotivasi, yang selalu menyirami do'a dan kasih sayang, serta menasehatiku menjadi lebih baik. Terima kasih atas semuanya semoga engkau di beri kesehatan dan panjang umur agar bisa menemani langkah kecilku menuju kesuksesan.



MOTTO

*“Apabila engkau telah selesai (dengan suatu urusan), teruslah bekerja keras
(untuk sesuatu yang lain)”*

(Q.S Al-Insyirah : 7)

*“Jangan pernah mau menunggu sampai kita siap, kerana kita tidak akan
menghabiskan sisa hidup kita hanya untuk menunggu ”*



KATA PENGANTAR

Segala puji syukur peneliti sampaikan kepada Allah SWT, karena atas Rahmat dan Hidayah-Nya, perencanaan, pelaksanaan dan penyelesaian tugas akhir/skripsi dengan judul “Implimentasi Data Mining Pada Penilaian Kinerja Guru Menggunakan Metode K-Nearest Neighbor (K-Nn) Studi Kasus Pada Mi Alfagiri Silo-Jember” sebagai salah satu syarat penyelesaian program diploma/sarjana dapat terselesaikan dengan baik dan lancar.

Kesuksesan ini dapat peneliti peroleh karena dukungan beberapa pihak. Peneliti menyampaikan terima kasih kepada :

1. KHR. Ach. Azaim Ibrahimi selaku Pengasuh Pondok Pesantren Salafiyah Syafi'iyah
2. KH. Ach Fadlail, S.H., M.H selaku Rektor Universitas Ibrahimi.
3. Abd. Ghofur, M.Kom selaku Dekan Fakultas Sains dan Teknologi.
4. Bapak Achmad Baijuri, M.Kom selaku Ketua Program Studi Sistem Informasi.A
5. Bapak Zaehol Fatah, M.Kom dan Bapak Achmad Baijuri, M.Kom selaku pembimbing I dan II, dan
6. Bapak M. Hosein F, S.Pd selaku kepala Madrasah Ibtida'iyah Alfagiri Krajan Pace Silo Jember.

Semoga semua amal baik yang telah diberikan oleh Bapak/Ibu kepada peneliti mendapat balasan yang sebaik mungkin dari Allah SWT, Amin.

Situbondo, ___ 2024

Muhammad Jundanuddin

DAFTAR ISI

HALAMAN JUDUL	i
PERNYATAAN KEASLIAN	Error! Bookmark not defined.
PERSETUJUAN PEMBIMBING.....	iii
HALAMAN PENGESAHAN	Error! Bookmark not defined.
HALAMAN PERSEMBAHAN.....	v
MOTTO	vi
KATA PENGANTAR.....	vii
DAFTAR ISI.....	vi
ABSTRAK	x
BAB I PENDAHULUAN	x
1.1 Latar Belakang.....	1
1.2 Identifikasi Masalah	4
1.3 Batasan Masalah.....	4
1.4 Tujuan Penelitian.....	4
1.5 Manfaat Penelitian.....	4
1.6 Sistematika Pembahasan	5
BAB II KAJIAN PUSTAKA	7
2.1 Penelitian Terdahulu.....	7
2.2 Landasan Teori	11
BAB III METODELOGI PENELITIAN.....	34
3.1 Metodologi Penelitian.....	34
3.1 Lokasi Penelitian	36
3.2 Bahan Penelitian.....	36
3.3 Metode Pengumpulan Data	36
3.4 Tahapan Penelitian	37
3.5 Pengolahan Data Mining	39
BAB IV HASIL DAN PEMBAHASAN.....	44
4.1 Gambaran Umum Lokasi Penelitian	44
4.2 Hasil.....	48
4.3 Pembahasan	49

BAB V PENUTUP	59
5.1 Kesimpulan.....	59
5.2 Saran.....	59
DAFTAR PUSTAKA	60
CURRICULUM VITAE	62



ABSTRAK

Muhammad Jundanuddin. 2024, Implimentasi Data Mining Pada Penilaian Kinerja Guru Menggunakan Metode K-Nearest Neighbor (K-NN) Studi Kasus Pada MI Alfagiri Silo-Jember. Pembimbing: (I) Zaehol Fatah, M.Kom., (II) Achmad Bajjuri, M.Kom.

MI Al fagiri Pace Silo Jember, Jawa Timur adalah sebuah Madrasah Ibtida'iyah yang terletak di Jember, Jawa Timur. MI Alfagiri merupakan salah satu lembaga swasta di Desa Pace, Kecamatan Silo yang setiap tahunnya menyelenggarakan Penilaian Kinerja Guru (PKG) untuk meningkatkan mutu pendidikan. Untuk itu MI Alfagiri merupakan lembaga yang senantiasa mendorong pengembangan profesionalisme guru dan meningkatkan kualitas tenaga pengajar Hal ini bertujuan untuk mengetahui kapasitas mengajar guru di sekolah tersebut. Evaluasi hasil mengajar guru tidak hanya untuk mengetahui kapasitas guru, namun juga merupakan bentuk evaluasi tahunan. Evaluasi hasil mengajar guru merupakan hal yang penting dalam rangka meningkatkan kualitas guru itu sendiri. Penilaian kinerja dapat dilakukan dengan menggunakan metode atau sistem supaya mendapatkan hasil yang lebih akurat dan lebih terperinci. Diantaranya yaitu Pengimplementasian data mining menggunakan metode K-Nearest Neighbor (KNN). Algoritma K-Nearest Neighbor (KNN) adalah algoritma machine learning yang bersifat non-parametric dan lazy learning. Dengan kata lain, tidak ada jumlah parameter atau estimasi parameter yang tetap dalam model, terlepas data tersebut berukuran kecil ataupun besar. Algoritma K-Nearest Neighbor (KNN) merupakan sebuah metode untuk melakukan klasifikasi terhadap objek berdasarkan data yang memiliki jarak paling dekat dengan objek tersebut. Tujuan dari algoritma k-nearest adalah untuk mengidentifikasi tetangga terdekat dari titik kueri tertentu sehingga kita dapat menetapkan label kelas pada titik tersebut. Algoritma ini hanya melakukan penyimpanan dan klasifikasi data. Misalnya pada tahap klasifikasi, fitur serupa dihitung untuk menguji data (mengklasifikasikan data yang tidak diketahui).

Kata kunci : *Penilaian Kinerja Guru, K-Nearest Neighbor (KNN)*

BAB I PENDAHULUAN

1.1 Latar Belakang

Guru merupakan pendidik yang mempunyai peran sangat penting dan tanggung jawab besar dalam pendidikan anak bangsa. Guru yang profesional diharapkan ikut serta dalam mencapai tujuan pendidikan nasional yang tertuang dalam Undang-Undang Nomor 20 Tahun 2003 tentang Sistem Pendidikan Nasional. Artinya menciptakan generasi bangsa Indonesia yang berakhlak mulia, berilmu dan terampil, serta bertaqwa kepada Tuhan Yang Maha Esa. Sehat jasmani, sehat rohani, berkepribadian stabil dan mandiri, serta rasa tanggung jawab sosial dan nasional. Guru merupakan orang yang mempunyai peranan penting dalam proses belajar mengajar. Secara umum guru adalah orang yang mengajar, memberi petunjuk, dan melatih peserta didik dalam pengetahuan dan keterampilan dalam berbagai bidang.

Tidak hanya memberikan pengetahuan tetapi juga membantu siswa mengembangkan keterampilan, sikap dan nilai-nilai yang penting bagi perkembangan pribadi dan sosialnya.

Kinerja guru merupakan suatu hal yang sangat penting dalam upaya meningkatkan semangat belajar para siswa di sekolah. Didalam dunia yang kompetitif dan menglobal. setiap lembaga, seperti Madrasah Ibtida'iyah Alfagiri memerlukan kinerja guru yang tinggi guna untuk meningkatkan kualitas Sumber Daya Manusia. Guru sebagai ujung tombak suatu lembaga memerlukan timbal balik dari lembaga atas hasil kerja yang mereka lakukan sebagai panduan bagi

perilaku mereka dimasa yang akan datang. timbal balik terhadap kinerja guru dapat dilakukan melalui evaluasi kinerja.

Penilaian kinerja guru penting dilakukan untuk perbaikan kualitas guru itu sendiri. Selain itu, penilain juga penting untuk lembaga pendidikan dalam merancang kembali dan cara baru untuk tercapainya tujuan pendidikan nasional. Penilain kinerja juga memberi masukan terhadap guru untuk mengevaluasi dan meningkatkan kinerja. Penilaian kinerja guru juga akan mempermudah guru dalam mengenal tugas-tugasnya secara lebih spesifik, sehingga bisa memberikan pembelajaran secara efektif dalam kemajuan siswa dan profesionalitas dalam mengajar.

Oleh sebab itu, penilaian kinerja guru penting dan perlu dilakukan oleh suatu sekolah untuk memperbaiki kinerja. Meski begitu, penting dipahami bahwa penilaian kinerja guru tidak dimaksudkan untuk mengkritik dan mencari kesalahan, melainkan sebagai dorongan bagi guru guna mengembangkan diri menjadi lebih professional, dan pada akhirnya akan meningkatkan kualitas pendidikan peserta didik. Untuk itu, perlu perubahan pola pikir dan perilaku serta kesediaan guru untuk refleksi diri secara berkelanjutan[1].

MI Alfabri merupakan lembaga swasta yang ada di desa pace, kecamatan silo, yang setiap tahunnya menyelenggarakan penilaian kinerja guru (PKG) dengan tujuan meningkatkan pendidikan yang bermutu, untuk itu, MI Alfabri adalah lembaga yang selalu mendorong peningkatan profesional guru dan meningkatkan kualitas tenaga pengajar, hal ini bertujuan untuk mengetahui kinerja guru di

lembaga tersebut. Penilaian Kinerja Guru bukan hanya untuk mengetahui kompetensi pengajar, akan tetapi sebagai bentuk evaluasi di setiap tahunnya.

Penilaian kinerja dapat dilakukan dengan menggunakan metode atau sistem supaya mendapatkan hasil yang lebih akurat dan lebih terperinci. Termasuk penerapan data mining dengan metode K-Nearest Neighbors (KNN). Algoritma K-Nearest Neighbors (KNN) adalah algoritma pembelajaran mesin yang belajar secara nonparametrik dan dengan penundaan. Metode nonparametrik berarti metode tersebut tidak membuat asumsi mengenai sebaran data yang mendasarinya. Dengan kata lain, tidak ada jumlah parameter atau estimasi parameter yang tetap dalam model, terlepas dari apakah datanya kecil atau besar.

. Algoritma K-Nearest Neighbor (KNN) merupakan sebuah metode untuk melakukan klasifikasi terhadap objek berdasarkan data yang memiliki jarak paling dekat dengan objek tersebut.

Tujuan dari algoritma k-nearest neighbor adalah untuk mengidentifikasi tetangga terdekat dari titik kueri yang diberikan, sehingga kita dapat menetapkan label kelas ke titik tersebut. Algoritma ini melakukan penyimpanan dan klasifikasi data. Contoh pada fase klasifikasi, fitur-fitur yang sama dihitung untuk testing data (klasifikasi data yang belum diketahui).[2]

Oleh karena itu, dengan latar belakang yang sudah dijelaskan diatas, maka sangat perlu kiranya menggunakan sistem dan metode, agar proses sistem berjalan dengan terorganisir. Maka dari itu peneliti berharap dengan adanya implementasi data mining pada penilaian kinerja guru menggunakan metode k-nearest neighbor (KNN) di MI Alfagiri Silo Jember dapat berjalan sebagaimana mestinya dan

membantu pihak lembaga dalam evaluasi guru dan mencari guru yang terbaik dalam melaksanakan kinerjanya.

1.2 Identifikasi Masalah

Berdasarkan latar belakang di atas, maka dapat dikemukakan permasalahan berikut ini.

1. Kurangnya ketepatan untuk menentukan penilaian pada kinerja para guru.
2. Perlunya metode agar keputusan yang diambil tepat sasaran

1.3 Batasan Masalah

Mengingat banyak aspek yang dijadikan pertimbangan dalam penyusunan laporan tugas akhir ini, maka perlu diberikan batasan masalah, yaitu:

1. Membangun sistem yang bisa menentukan nilai kinerja guru dengan tepat menggunakan Metode K-Nearest Neighbor (KNN).
2. Penerapan metode K-Nearest Neighbor (KNN) ini untuk melakukan penghitungan kinerja guru di MI Alfagiri.

1.4 Tujuan Penelitian

Adapun tujuan yang ingin dicapai dari penelitian ini adalah mengimplementasikan data mining pada penilaian kinerja guru menggunakan metode K-Nearest Neighbor (KNN). Agar memberikan masukan bagi guru untuk memperbaiki dan meningkatkan kinerjanya.

1.5 Manfaat Penelitian

Adapun manfaat yang diharapkan dari penelitian ini ialah:

1. Membantu dan mempermudah pihak sekolah untuk penilaian kinerja para guru.

2. Memberikan tambahan informasi bagi sekolah mengenai potensi kinerja guru.
3. Mengetahui Guru yang terbaik kinerjanya.

1.6 Sistematika Pembahasan

Pembahasan sistematis ini dimaksudkan agar mudah dibaca dan komprehensif guna memberikan pemahaman yang utuh dan akurat terhadap isi penelitian ini. Bab pertama memuat latar belakang masalah, identifikasi masalah dan rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, metode penelitian, dan sub metode penelitian meliputi jenis penelitian, pengumpulan data. Meliputi metode, metode pengembangan sistem, dan diskusi sistematis di bagian akhir. Secara umum bab ini berharap pembaca dapat menemukan latar belakang teoritis dan dasar pemikiran dari sumber terpercaya dan situasi realistis di bidang penelitian. Lebih lanjut, bab ini juga membahas kedudukan laporan dalam bidang keilmuannya masing-masing, dengan tetap menjaga kesinambungan dengan keilmuan sebelumnya. Dapat disimpulkan bahwa bab ini merupakan landasan atau acuan metodologis untuk bab-bab selanjutnya. Artinya isi bab selanjutnya adalah pengembangan teori. Hal ini lebih terfokus pada pendukung atau penguatan teori-teori yang berdasarkan atau diacu pada Bab 1, bukan sebagai tolak ukur pengembangan teori.

BAB I : PENDAHULUAN

Pada bab ini menguraikan tentang latar belakang masalah, identifikasi masalah, rumusan masalah, landasan masalah, tujuan dan manfaat, metode pengumpulan data, metode pengembangan sistem, serta sistematika pembahasan.

BAB II : KAJIAN PUSTAKA

pada bab ini dipaparkan tentang tinjauan pustaka penunjang, hal – hal yang terkait dengan system serta referensi panunjang yang sesuai dengan judul.

BAB III : ANALISIS DAN PERANCANGAN SISTEM

Bab ini memaparkan tentang analisis dan desain. Memberikan gambaran tentang obyek penelitian, alur proses, desain sistem, desain proses, desain database dan desain interface.

BAB IV : IMPLEMENTASI SISTEM

Pada bab ini di paparkan tentang kontruksi sistem yang berisi kebutuhan sistem, instalasi sistem, source code, serta cara kerja sistem.

BAB V : PENUTUP

Pada bab ini dibahas kesimpulan dari laporan yang dibuat serta saran untuk pengembangan aplikasi, lembaga maupun untuk instansi.

BAB II

KAJIAN PUSTAKA

2.1 Penelitian Terdahulu

Dalam penelitian ini, penulis mengacu kepada penelitian-penelitian yang pernah dilakukan oleh peneliti-peneliti terdahulu sebagai referensi diantaranya sebagai berikut:

2.1.1 Penerapan Metode K-Nearest Neighbor Untuk Klasifikasi Kinerja Satpam Berbasis Web[3]

Penilaian kinerja adalah salah satu aspek penting untuk mengembangkan sebuah organisasi atau instansi agar menjadi lebih efektif dan efisien. Dimana penilaian kinerja ini dapat dijadikan sebagai tolak ukur dalam mengambil keputusan. Sedangkan dengan tidak adanya penilaian kinerja, maka kualitas kinerja para satpam menjadi tidak terkontrol. Satpam dengan kinerja yang baik tidak akan mendapatkan penghargaan atas kinerjanya. Begitu pula sebaliknya, satpam dengan kinerja yang buruk juga tidak mendapatkan sanksi apapun. Keadaan ini akan berdampak pada menurunnya motivasi satpam untuk meningkatkan kualitas kinerja mereka dan cenderung menurun. Dengan menurunnya kinerja satpam maka akan berdampak pada menurunnya tingkat keamanan di sebuah instansi. Peneliti akan membangun sistem untuk mengklasifikasikan kinerja satpam agar menjadi referensi bagi instansi untuk bisa menindak lanjuti satpam yang memiliki kinerja yang buruk serta memberikan apresiasi bagi satpam yang kinerjanya baik. Klasifikasi tersebut dikerjakan menggunakan metode K-Nearest Neighbor untuk menghitung

tingkat akurasi dari penilaian kinerja satpam. Penelitian ini akan terdapat data sebanyak 130 data dengan 10 atribut, yang berisikan tentang penilaian kinerja seperti kepribadian, kemampuan dan keterampilan. Hasil penelitian akan menghasilkan output berupa penilaian tingkat kinerja dari satpam termasuk kategori baik, cukup atau kurang.

Dari hasil analisa dan perancangan sistem klasifikasi kinerja satpam dengan menggunakan metode k-nn, maka dapat diambil kesimpulan Proses penerapan metode K-Nearest Neighbor untuk klasifikasi kinerja satpam menggunakan perhitungan K=3 sampai K=9. Pengujian klasifikasi menggunakan 6 data uji dan menghasilkan prosentase terbesar dengan k=3 yaitu akurasi sebesar 66,67%. Dengan demikian aplikasi ini bisa digunakan untuk membantu mengetahui kinerja satpam.

2.1.2 Implementasi Algoritma K-Nearest Neighbor (Knn) Untuk Memprediksi Kinerja Siswa[4]

Salah satu unsur untuk menjadi penilaian akreditasi adalah tepat waktu lulusan siswa. Adanya siswa yang tidak aktif tentu akan mempengaruhi panjangnya waktu kelulusan. Prediksi kinerja siswa diperlukan untuk mencegah siswa yang tidak aktif. Algoritma KNN digunakan untuk memprediksi kinerja siswa dengan menggunakan metode klasifikasi. Penelitian ini untuk mengoptimalkan algoritma KNN untuk memprediksi kinerja siswa dengan metode klasifikasi. Penelitian yang telah dilakukan dengan menggunakan data Jurusan Teknik Informatika Politeknik Harapan Bersama menyimpulkan bahwa nilai K terbaik adalah 3, 6, dan 9 untuk mendapatkan prediksi terbaik. Hasil ini

diperoleh dengan mencoba nilai K, 3 hingga 60. Nilai prediksi kemudian dibandingkan, hasil yang salah diprediksi dimana persentase terkecil adalah yang terbaik

Penelitian yang dilakukan dengan menggunakan data Departemen Teknologi Informasi Universitas Teknologi Harapan Bersama menyimpulkan bahwa nilai K optimal untuk memperoleh prediksi terbaik adalah 3, 6, dan 9.

Hasil ini diperoleh dengan melakukan percobaan dengan nilai K dari 3 hingga 60. Nilai prediksi tersebut kemudian dibandingkan dan nilai dengan persentase hasil prediksi salah terendah adalah yang terbaik.

2.1.3 Implementasi Algoritma Knn Untuk Memprediksi Performa Siswa

Sekolah[5]

Salah satu faktor yang mempengaruhi tingkat kelulusan siswa adalah prestasi belajarnya. Memprediksi tingkat kelulusan berdasarkan kinerja siswa mempunyai manfaat untuk menganalisis siswa yang kinerjanya buruk secara akademis dan memberikan dukungan kepada siswa yang menghadapi kesulitan dalam proses pembelajaran. Ada beberapa faktor yang perlu diperhatikan dalam memprediksi tingkat kelulusan siswa, seperti nilai akademik, sikap, dan faktor sosial. Namun, faktor-faktor ini saja tidak cukup untuk memprediksi kinerja siswa secara efektif, dan pendidik juga kesulitan mengidentifikasi faktor mana yang mempengaruhi kinerja siswa. Untuk memprediksi kinerja siswa sekolah, digunakan metode K-Nearest Neighbor (KNN). Metode K-Nearest Neighbor sering digunakan dalam mengklasifikasikan kinerja siswa karena kesederhanaannya dan kemampuannya dalam menghasilkan hasil yang

signifikan dan kompetitif. Dalam penelitian ini prediksi tingkat kelulusan siswa dilakukan dengan menggunakan metode KNN. Hasil penerapan prediksi kinerja siswa dengan metode KNN dapat menjadi acuan siswa untuk meningkatkan prestasinya dan membantu pendidik dalam mempertimbangkan masa depan. bahan ajar.

Untuk dapat menghasilkan SDM (sumber daya manusia) yang cakap, berwawasan, kompetitif dan kreatif, lembaga pendidikan diminta untuk menyelenggarakan pendidikan yang bermutu dan berkualitas bagi peserta didiknya. Sehingga untuk mengimplementasikannya dilakukan beberapa perubahan kurikulum guna meningkatkan kualitas pada tingkatan pendidikan tentunya untuk mengetahui berhasil atau tidaknya kurikulum yang telah dibuat diperlukan sebuah evaluasi berupa prediksi hasil belajar siswa di sekolah serta faktor-faktor yang mempengaruhinya.

Maka untuk mewujudkan hal tersebut dibutuhkan suatu alat analisis berupa komputasi cerdas untuk menganalisis bagaimana kinerja siswa, faktor mana yang akan mempengaruhi kinerja mereka, dengan cara apa siswa dapat membuat kemajuan, dan apakah siswa memiliki potensi untuk tampil lebih baik.

Pada penelitian ini akan dilakukan klasifikasi performa siswa berdasarkan lingkungan serta hasil akademik siswa dengan K-Nearest Neighbor (KNN) algoritma ini digunakan karena kemampuannya dalam menyederhanakan perhitungan algoritma dan mengoptimalkan waktu. KNN adalah metode yang bekerja dengan mengelompokkan data baru berdasarkan jarak mereka dengan

data lainnya. Prediksi akan dilakukan menggunakan algoritma K-Nearest Neighbor (KNN).

Penelitian ini menyimpulkan bahwa telah berhasil dibuat sebuah aplikasi untuk melakukan seleksi data penerima beasiswa dengan menggunakan algoritma KNN. Evaluasi algoritma KNN menggunakan metode confusion matrix menunjukkan hasil rata-rata akurasi dari metode KNN ini sebesar 96%. Ini menunjukkan bahwa implementasi algoritma KNN pada prediksi performa siswa sekolah memiliki akurasi yang cukup tinggi dengan demikian klasifikasi performa dengan KNN dapat dijadikan sebagai system pendukung untuk membantu guru serta instansi pendidikan untuk mempertimbangkan langkah-langkah yang dapat dilakukan untuk pemajuan pendidikan di Indonesia kedepannya.

2.2 Landasan Teori

2.2.1 Data

Data adalah bahan penyusun dasar informasi yang kemudian diproses lebih lanjut agar menghasilkan informasi. Data juga dapat di artikan suatu ide, objek, atau angka, karakter, gambar, atau simbol apa pun yang dapat dimasukkan (input) atau dikeluarkan (output) oleh komputer.

Data adalah dasar informasi. Setelah data diproses dan dianalisis, wawasan yang bermakna dapat diperoleh dan pengambilan keputusan dapat didukung. Dalam konteks teknologi informasi, data seringkali diatur dan disimpan dalam format tertentu sehingga komputer dapat dengan mudah mengakses dan memprosesnya.

2.2.2 Basis Data

Basis data memiliki beberapa pengertian dari sudut pandang yang berbeda di antaranya:

- a. Kumpulan kelompok data yang saling berhubungan yang diorganisasi sedemikian rupa agar kelak dapat dimanfaatkan kembali dengan cepat dan mudah.
- b. Kumpulan data yang saling berhubungan yang disimpan secara bersama sedemikian rupa dan tanpa pengulangan (redudansi) yang tidak perlu, untuk memenuhi berbagai kebutuhan
- c. Kumpulan file/tabel yang saling berhubungan yang disimpan dalam media penyimpanan elektronik.

Menurut beberapa definisi di atas, basis data adalah kelompok data yang saling berhubungan dan terorganisir menurut aturan tertentu sehingga dapat digunakan untuk berbagai tujuan dan disimpan dalam media penyimpanan elektronik. Basis data juga dapat didefinisikan menjadi media penyimpanan data untuk dapat diakses dengan mudah dan cepat.[6] Sedangkan menurut pendapat yang lain menerangkan bahwa Database merupakan sekumpulan tabel-tabel yang berisi data dan merupakan kumpulan dari field atau kolom. Struktur file yang menyusun sebuah database adalah Data Record dan Field. Jadi basis data adalah media untuk menyimpan data yang mana merupakan tabel-tabel yang berisi data dan merupakan kumpulan dari field dan kolom.

Tujuan utama dari pengelolaan data dalam sebuah basis data adalah agar kita dapat menemukan kembali data (yang kita cari) dengan mudah dan cepat, tujuan basis data di bedakan menjadi 2 kelompok, yaitu :

- a. Tujuan Primer, tujuan utama yang ingin dicapai dalam usaha perancangan dan pengembangan basis data
- b. Tujuan Sekunder, tujuan tambahan yang di maksudkan untuk mencapai tujuan primer.

2.2.3 Data Mining

Data mining ialah proses mengekstraksi pengetahuan dan informasi berharga dari kumpulan data yang besar dan kompleks. Tujuan utama dari data mining adalah mengidentifikasi pola, hubungan, atau informasi yang mungkin tidak terlihat secara langsung dalam data, sehingga dapat memberikan wawasan yang lebih dalam dan bernilai.[7]

Proses data mining melibatkan penggunaan berbagai teknik statistik, matematis, dan kecerdasan buatan untuk menganalisis data dengan cara yang sistematis dan otomatis. Hasil dari data mining bisa digunakan dalam mendukung penilaian kinerja, digunakan untuk memahami tren pasar, meningkatkan efisiensi bisnis atau merumuskan strategi bisnis. Berikut adalah beberapa definisi Data Mining menurut para ahli:

- a. Menurut Han dan Kamber: “Data mining adalah proses penemuan pola atau informasi yang berguna dari basis data besar dengan

- menggunakan metode termasuk teknik statistik, matematika, dan kecerdasan buatan.”
- b. Menurut Berry dan Linoff: “Data mining adalah proses penemuan pola berharga atau informasi pengetahuan baru dalam database besar dengan menggunakan algoritma pencarian atau algoritma pembelajaran mesin.”
 - c. Menurut Fayyad, Piatetsky-Shapiro, dan Smyth: “Data mining adalah proses ekstraksi pengetahuan yang berharga, pola, atau informasi tersembunyi dari sejumlah besar data”.
 - d. Margaret H. Dunham: “Data mining adalah proses otomatisasi dalam penemuan pola yang bermakna, kecocokan, korelasi, dan tren yang tersembunyi dalam database.”
 - e. Usama Fayyad: “Data mining ialah pemrosesan penggalian pengetahuan yang dimaksudkan untuk menemukan pola-pola yang bermanfaat, pengetahuan baru, dan informasi yang tersembunyi dari data yang tersimpan di basis data”. [8]

2.2.4 Peran data Mining

Ada beberapa peran utama dari data mining. Berikut adalah peran utama data mining :



Gambar 2.1 Peran Data Mining

1. Estimasi

Dalam estimasi, satu set variabel prediktor numerik dan/atau kategoris digunakan untuk mendekati nilai variabel target numerik. Model dibangun menggunakan dataset “lengkap” yang memberikan nilai untuk variabel target dan prediktor. Kemudian untuk observasi baru, nilai variabel target diperkirakan berdasarkan nilai variabel prediktor. Misalnya, mengestimasi biaya belanja keperluan sekolah untuk empat keluarga yang dipilih secara acak musim gugur ini. Contoh algoritma untuk mengimplementasikan estimasi adalah neural networks

2. Prediksi

Prediksi mirip dengan klasifikasi dan estimasi, perbedaannya terletak pada prediksi, dan hasilnya ada di masa depan. Misalnya, memprediksi harga saham dalam 3 bulan ke depan, atau jika batas kecepatan dinaikkan, memprediksi persentase peningkatan kematian lalu lintas tahun depan. Jika sesuai, setiap metode dan teknik yang digunakan untuk klasifikasi dan estimasi juga dapat digunakan Untuk prediksi. Hal ini termasuk teknik statistik tradisional seperti estimasi titik dan estimasi interval kepercayaan, regresi dan korelasi linier sederhana, dan regresi berganda.. Contoh algoritma dalam mengimplementasikan prediksi adalah k-nearest neighbor, decision tree, dan neural networks.

3. Klasifikasi

Klasifikasi mirip dengan estimasi, dengan pengecualian bahwa variabel objektif bersifat kategoris daripada numerik. Ada variabel kategoris tujuan dalam klasifikasi, contohnya kelompok pendapatan, yang dapat dibagi menjadi tiga kelompok atau kategori: pendapatan tinggi, pendapatan menengah, dan pendapatan rendah. Model data mining mengevaluasi sejumlah besar catatan, yang masing-masing berisi informasi tentang variabel prediktor. Algoritma akan bekerja kira-kira dengan cara berikut. Periksa kumpulan data yang mencakup variabel prediktor dan variabel target (yang sudah dikategorikan), kelompok pendapatan. Dalam pendekatan ini, algoritma (perangkat lunak) "mempelajari" variabel mana yang terhubung dengan rentang pendapatan tertentu. Misalnya ada seorang wanita tua mungkin berkorelasi dengan kategori berpenghasilan tinggi. Pengumpulan data ini disebut sebagai training set.

Selanjutnya, algoritma akan melihat entri baru yang tidak tersedia informasi braket pendapatan. Algoritma akan memberikan kategori ke record baru berdasarkan klasifikasi dalam set pelatihan. Contoh algoritma untuk mengimplementasikan klasifikasi adalah k-nearest neighbor, decision tree, dan neural networks.

4. Klastering

Pengelompokan catatan, pengamatan, atau kejadian ke dalam kelas item terkait disebut sebagai klastering. Klastering adalah sekelompok record yang dapat dibandingkan satu sama lain tetapi tidak

dengan record di klaster lain. Klastering bervariasi dari klasifikasi karena tidak memiliki variabel target. Tugas klastering tidak berusaha untuk mengkategorikan, memperkirakan, atau memperkirakan nilai variabel target. Algoritma klastering, di sisi lain, mencoba untuk membagi seluruh kumpulan data menjadi subkelompok atau klaster yang relatif homogen, di mana kesamaan catatan di dalam klaster dimaksimalkan dan kesamaan dengan catatan di luar klaster ini diminimalkan. Target pemasaran produk khusus untuk perusahaan kecil dengan anggaran pemasaran terbatas atau untuk membagi aktivitas keuangan menjadi kelompok yang tidak berbahaya dan mencurigakan untuk alasan audit akuntansi adalah contoh dari klastering. Klastering pun sering digunakan sebagai tahap pertama dalam proses data mining, dengan klaster yang dihasilkan digunakan sebagai masukan lebih lanjut ke dalam pendekatan yang berbeda di kemudian hari. Contoh algoritma untuk mengimplementasikan klastering adalah k-means.

5. Asosiasi

Upaya menentukan kualitas mana yang "berjalan bersama" dalam data mining dikenal sebagai tugas asosiasi. Pekerjaan asosiasi mencoba menemukan kriteria untuk mengukur hubungan antara dua atau lebih kualitas. Ini paling umum di sektor korporasi, yang dikenal sebagai analisis afinitas atau analisis keranjang pasar. Aturan asosiasi berbentuk "Jika anteseden maka konsekuen," bersama dengan ukuran dukungan dan kepercayaan aturan.

Contoh paling umum adalah, jika 10 konsumen berbelanja di supermarket dan membeli roti, jika 8 dari 10 orang konsumen membeli roti dengan selai maka aturan asosiasinya adalah “jika membeli roti, maka belilah selai”. Contoh algoritma untuk mengimplementasikan asosiasi adalah algoritma apriori dan algoritma GRI.

2.2.5 Masalah dalam Data Mining

Masalah dalam data mining Sistem basis data adalah sistem (konfigurasi) yang digunakan untuk mendukung proses tertentu. Sistem database terdiri dari sekumpulan tabel yang melayani suatu tujuan. Misalnya, untuk membuat sistem pencatatan pasien, Anda memerlukan tiga tabel dasar: tabel biodata pasien, tabel data penyakit, dan tabel pengobatan. Sistem basis data adalah objek pasif. Sistem ini hanya berfungsi sebagai penyimpanan, sehingga pengguna tidak dapat menggunakan sistem dari jarak jauh.

Operasi yang biasa dilakukan dalam sistem basis data adalah membuat, memperbarui, dan menghapus. Sistem database biasanya disimpan di server.

Lebih tepatnya, sistem database terdiri dari beberapa komponen.

- Hardware
- Sistem Operasi
- Basis data
- Sistem
- Pemakai
- Aplikasi (Perangkat lunak yang digunakan untuk mengoptimalkan database)

Permasalahan lain pada sistem database adalah ketika pertama kali dikembangkan hanya berfungsi sebagai alat pendukung bisnis untuk mengurangi jumlah pekerjaan. Tidak ada tujuan yang jelas bagi sistem basis data untuk memproses dan menghasilkan pengetahuan.

Untuk mempersiapkan sistem database Anda untuk mencari pengetahuan, Anda harus melakukan beberapa langkah proses Pembersihan, normalisasi, dan peringkasan. Selain itu, Hilangnya data pada suatu database dapat terjadi karena berbagai faktor seperti kerusakan hardware, software, serangan malware, dan human error.

Anda dapat mengambil beberapa langkah untuk mengatasi masalah ini.

- Pencadangan data rutin di berbagai media penyimpanan (seperti penyimpanan cloud atau perangkat eksternal) merupakan langkah penting untuk memulihkan data yang hilang. Pencadangan harus dijadwalkan pada frekuensi yang memenuhi kebutuhan bisnis Anda.
- Menerapkan Teknik redundansi seperti RAID (redundant array of Independent Disks) dan replikasi data ke server lain dapat membantu mencegah kehilangan data akibat kerusakan fisik pada perangkat keras.
- Menggunakan transaksi dalam pemrosesan ACID (Atomicity, Consistency, Isolation, Durability) adalah seperangkat prinsip yang harus diterapkan saat mengelola transaksi database untuk

memastikan data tetap konsisten, sehingga Anda terlindungi dari kerugian selama transaksi proses.

- Log dan jejak audit. Menyimpan log perubahan dan jejak audit aktivitas dalam database membantu pelacakan kesalahan dan memungkinkan data dipulihkan dari titik waktu tertentu jika terjadi masalah.
- Menerapkan kontrol akses dan keamanan. Mengelola hak akses pengguna dengan benar dan menerapkan sistem keamanan yang kuat (enkripsi, firewall, deteksi intrusi, dll.) untuk mencegah serangan dunia maya yang dapat menyebabkan kehilangan data
- Pemulihan otomatis: Beberapa sistem manajemen basis data (DBMS) menyediakan fitur pemulihan otomatis yang dapat memulihkan basis data ke kondisi terakhir yang konsisten setelah kegagalan atau kerusakan.
- Menerapkan prosedur validasi dan validasi data selama input dapat membantu mencegah kesalahan pada data yang belum diproses atau salah.
- Pemantauan dan pemeliharaan rutin terhadap database dan infrastruktur terkait membantu mengidentifikasi potensi masalah sebelum menyebabkan kehilangan data.

Dengan menggabungkan strategi di atas, Anda dapat meminimalkan risiko kehilangan data dan memulihkan data yang hilang dengan lebih mudah.

Data yang noise cenderung sulit diproses menggunakan teknik statistik seperti Gaussian. Untuk mengolah data secara non-aural harus dilakukan proses pembersihan.

2.2.6 Prediksi

Prediksi adalah suatu proses memperkirakan secara sistematis tentang sesuatu yang paling mungkin terjadi di masa mendatang berdasarkan informasi masa lalu dan sekarang yang dimiliki, agar kesalahannya dapat terminimalisir. Prediksi tidak harus memberikan jawaban secara pasti kejadian yang akan terjadi, melainkan berusaha untuk mencari jawaban sedekat mungkin yang akan terjadi.

Pengertian Prediksi sama dengan ramalan atau perkiraan. Menurut kamus besar bahasa Indonesia, prediksi adalah hasil dari kegiatan memprediksi atau meramal atau memperkirakan nilai pada masa yang akan datang dengan menggunakan data masa lalu. Prediksi menunjukkan apa yang akan terjadi pada suatu keadaan tertentu dan merupakan input bagi proses perencanaan dan pengambilan keputusan. Prediksi bisa berdasarkan metode ilmiah ataupun subjektif belaka. Ambil contoh, prediksi cuaca selalu berdasarkan data dan informasi terbaru yang didasarkan pengamatan termasuk oleh satelit. Begitupun prediksi gempa, gunung meletus ataupun bencana secara umum. Namun, prediksi seperti pertandingan sepakbola, olahraga, dll umumnya berdasarkan pandangan subjektif dengan sudut pandang sendiri yang memprediksinya.

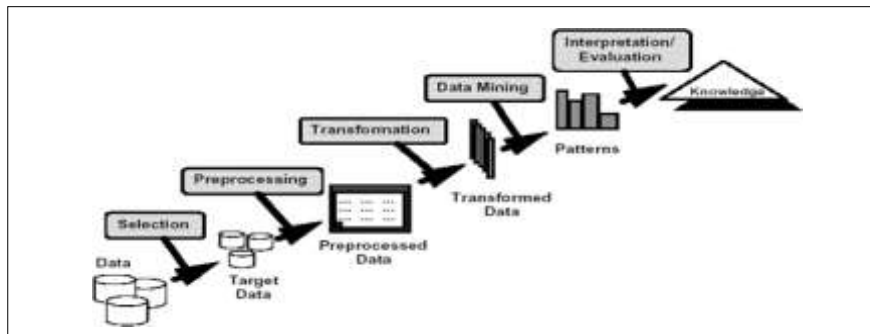
Berikut inilah beberapa pengertian prediksi atau peramalan menurut para ahli:

1. Prediksi (forecasting) adalah seni dan ilmu yang memprediksi peristiwa masa depan. Peramalan memerlukan pengambilan data historis dan menerapkan ke masa depan (Heizer & Render, 2011).
2. Prediksi adalah pemikiran terhadap suatu besaran, misalnya permintaan terhadap satu atau beberapa produk pada periode yang akan datang (Ishak, 2010).
3. Prediksi adalah memperkirakan beberapa kebutuhan pada periode yang datang meliputi kebutuhan dalam ukuran kuantitas (jumlah), kualitas (mutu), waktu dan lokasi yang dibutuhkan dalam rangka memenuhi permintaan barang atau jasa (Nasution, 2003).[9]

2.2.7 Knowledge Discovery in Database

Knowledge Discovery in Database (KDD) adalah proses menemukan informasi dan pola yang berguna dalam data. Informasi ini terkandung dalam database yang besar, yang sebelumnya tidak diketahui, dan berpotensi berguna.. Knowledge Discovery in Database (KDD) adalah kegiatan yang meliputi pengumpulan, pemakaian data, historis untuk menemukan keteraturan, pola atau hubungan dalam set data berukuran besar[10]. Data Mining merupakan salah satu langkah dari serangkaian proses iterative KDD. Berdasarkan pengertian tersebut dapat ditarik kesimpulan bahwa Knowledge Discovery in Database (KDD) adalah proses yang bertujuan untuk menggali dan menganalisis sejumlah besar himpunan data dan mengekstrak informasi serta pengetahuan yang berguna.

Langkah penting dalam proses KDD dapat dilihat pada gambar 2.1 yang terdiri dari tahapan-tahapan sebagai berikut.



Gambar 2.1 Tahapan Dalam KDD

Tahapan proses KDD adalah sebagai berikut:

1. Pembersihan Data: Fase ini membersihkan atau menyaring data yang tidak lengkap, tidak akurat, atau tidak konsisten. Hal ini penting untuk menghindari noise dan memastikan kualitas data yang digunakan untuk analisis.
2. Integrasi Data: Gabungkan data dari berbagai sumber ke dalam kumpulan data yang konsisten. Data dapat berasal dari beberapa database, file, atau gudang data dan harus diintegrasikan untuk analisis lebih lanjut.
3. Pemilihan data: Pada fase ini, data yang relevan dengan proses penambangan diidentifikasi dan dipilih dari kumpulan data terintegrasi. Data yang tidak relevan atau tidak berguna akan diabaikan atau dihapus.

4. Transformasi Data: Data yang dipilih diubah atau diubah ke dalam format yang sesuai untuk data mining. Proses ini mungkin mencakup normalisasi, agregasi, atau pengkodean ulang atribut untuk memfasilitasi analisis.
5. Data Mining: Ini adalah fase inti dari proses KDD di mana metode dan algoritma data mining diterapkan untuk menemukan pola, tren, atau informasi berguna dari data. Teknik yang digunakan antara lain klasifikasi, clustering, asosiasi, dan regresi, tergantung tujuan analisisnya.
6. Evaluasi Pola: Setelah suatu pola ditemukan, fase ini melibatkan evaluasi pola untuk menentukan apakah pola tersebut bermakna dan berguna. Pola yang relevan dan berguna dipilih dan pola yang kurang relevan diabaikan.
7. Presentasi Pengetahuan: Hasil proses KDD disajikan dalam format yang mudah digunakan. Hal ini dapat berupa visualisasi data, laporan, grafik, atau format lainnya yang memungkinkan pengguna membuat keputusan yang tepat.

2.2.6 Klasifikasi

Klasifikasi adalah proses mempelajari fungsi tujuan (objektif) f yang menugaskan setiap kumpulan atribut x ke salah satu label kelas yang telah ditentukan sebelumnya. Fungsi tujuan disebut juga model klasifikasi. Klasifikasi adalah menemukan model yang menjelaskan atau membedakan konsep atau

kelas dalam data, dengan tujuan untuk dapat menyimpulkan kelas suatu objek yang kelasnya tidak diketahui..

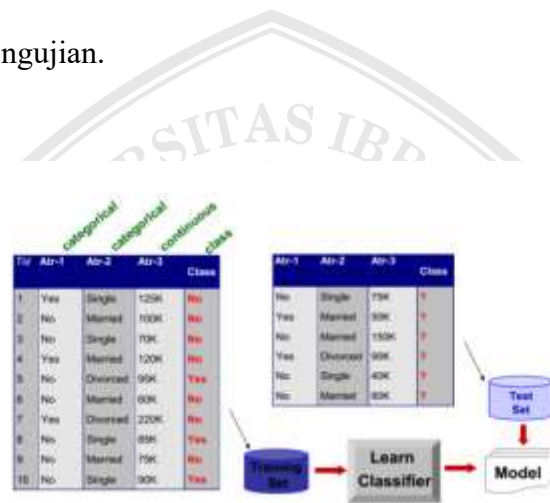
Dalam klasifikasi, terdapat target variabel kategori. Sebagai contoh, penggolongan pendapatan dapat dipisahkan dalam tiga kategori, yaitu pendapatan tinggi, pendapatan sedang, dan pendapatan rendah. Memperkirakan apakah suatu pengajuan hipotek oleh nasabah merupakan suatu kredit yang baik atau buruk[11]. Di dalam klasifikasi diberikan sejumlah record yang dinamakan training set, yang terdiri dari beberapa atribut, atribut dapat berupa kontinyu ataupun kategoris, salah satu atribut menunjukkan kelas untuk record.

Klasifikasi terdiri dari dua jenis model, yaitu:

1. Pemodelan deskriptif, merupakan model klasifikasi yang berfungsi sebagai suatu alat penjelasan untuk membedakan objek-objek dalam kelas-kelas yang berbeda.
2. Pemodelan prediktif, merupakan klasifikasi yang dapat digunakan untuk memprediksi label kelas record yang tidak ketahui.

Klasifikasi adalah tugas data mining yang memetakan data ke dalam kelompok-kelompok kelas. Teknik klasifikasi melakukan pengklasifikasian item data ke label kelas yang telah ditetapkan, membangun model klasifikasi dari kumpulan data input, membangun model yang digunakan untuk memprediksi tren data masa depan. Algoritma yang umum digunakan meliputi K-Nearest neighbor, Naïve Bayes Classification, Pohon Keputusan, Neural Network, dan Suport Vector Machines.

Klasifikasi digunakan untuk menentukan kumpulan data baru ke salah satu data yang telah ditentukan sebelumnya. Kelas ini disebut data pelatihan. Metode yang menggunakan klasifikasi data antara lain pohon keputusan, K-nearest neighbours, dan Naive Bayes. Klasifikasi data melibatkan perbandingan data baru (data uji) dengan data pelatihan. Hasil perbandingan ini menghasilkan keputusan mengenai hasil data pelatihan yang diinginkan. Di bawah ini adalah contoh tabel yang digunakan untuk membandingkan data pelatihan dan pengujian.



Gambar 2.1 Data Pelatihan dan Data Uji

Data pelatihan adalah data masa lalu. Tujuan dari data ini adalah sebagai sumber pengetahuan yang akan muncul kembali di masa depan. Oleh karena itu, jika beberapa karakter sudah disimpan dalam data pelatihan dan data pengujian dimanipulasi agar datanya cocok dengan data pelatihan, data pengujian mungkin memiliki properti yang sama dengan data pelatihan. Semakin besar jumlah data pelatihan, semakin akurat keputusan yang diambil melalui data mining. Klasifikasi juga dapat meningkatkan kemampuan Anda

dalam melakukan sesuatu dengan meningkatkan data pelatihan Anda. Kemampuan belajar dari metode klasifikasi disebut learning.

2.2.7 K-Nearest Neighbor

Metode K-Nearest Neighbor merupakan suatu metode untuk klasifikasi, metode ini melakukan klasifikasi terhadap suatu obyek yang berdasarkan kepada data training yang mempunyai jarak yang paling dekat dari obyek tertentu.

K-nearest neighbours merupakan algoritma klasifikasi yang menentukan kategori berdasarkan kategori terbanyak dari K-nearest neighbours. Hal ini dilakukan dengan mencari kelompok k objek terdekat pada data latih (Yang paling serupa), serupa dengan objek pada data baru atau data uji. K-Nearest Neighbor adalah metode mengklasifikasikan objek baru berdasarkan (K) tetangga terdekatnya.

K-NN adalah algoritma pembelajaran terbimbing yang mengklasifikasikan hasil query instance baru berdasarkan sebagian besar kategori di K-NN. Kategori yang paling umum adalah kelas hasil klasifikasi.

Metode atau algoritma K-Nearest Neighbor (K-NN) adalah metode untuk klasifikasi, metode K-NN melakukan klasifikasi suatu obyek yang berdasar kepada data training yang mempunyai jarak yang paling dekat dari obyek tersebut[9].

Metode K-Nearest Neighbor (K-NN) menggunakan algoritma terawasi dimana hasil query instance baru diklasifikasikan ke dalam K tetangga terdekat berdasarkan kategori terbanyak. Tujuan dari algoritma ini adalah untuk

mengklasifikasikan objek baru berdasarkan atribut dan contoh pelatihannya. Pada algoritma K-Nearest Neighbor (KNN) terdapat 5 jarak atau cara untuk mencari tetangga terdekat yaitu:

1. Euclidean Distance
2. Manhattan Distance
3. Cosine Distance
4. Corelation Distance
5. Hamming Distance

Dalam penelitian ini menggunakan cara Euclidean Distance untuk mencari tetangga terdekat, dengan rumus sebagai berikut :

$$(x, y) = \sqrt{\sum p (xi - yi)^2}$$

Nilai Xi merupakan nilai yang ada pada data training, sedangkan nilai Yi merupakan nilai yang ada pada data testing, nilai K merupakan dimensi atribut.

Langkah-langkah untuk menghitung algoritma K-NN

1. Menentukan nilai K
2. Menghitung kuadrat jarak euclid (query instance) masing masing objek terhadap training data yang diberikan
3. Kemudian mengurutkan objek-objek tersebut kedalam kelompok yang mempunyai jarak euclid terkecil
4. Mengumpulkan label class Y (klasifikasi Nearest Neighbor)

5. Dengan menggunakan kategori Nearest Neighbor yang paling mayoritas maka dapat diprediksikan nilai Query Instance yang telah dihitung.

KNN adalah algoritma pembelajaran terawasi yang mengklasifikasikan hasil instance kueri baru berdasarkan sebagian besar kategori di KNN. Kelas yang paling umum adalah kelas hasil klasifikasi. Algoritme K-NN kuat terhadap data pelatihan yang noise dan merupakan metode klasifikasi yang efektif untuk kumpulan data besar. Algoritme K-NN sangat konsisten; bahkan ketika jumlah data mendekati tak terhingga, algoritme tersebut menjamin tingkat kesalahan kurang dari dua kali tingkat kesalahan Bayesian (tingkat kesalahan minimum untuk distribusi data tertentu).[12]

2.2.8 Konsep K-Nearest Neighbor

K-Nearest Neighbors (K-NN) adalah kumpulan dari 10 teknik data mining yang terpopuler. Metode K-NN murni masuk dalam kategori pembelajar yang malas karena menunda proses pelatihan (atau tidak melakukan pelatihan sama sekali) hingga ada data uji yang label kelasnya ingin diketahui.

Algoritma KNN melakukan klasifikasi berdasarkan kemiripan suatu data dengan data lainnya.

2.2.9 Rapidminer

Rapid Miner adalah perangkat lunak yang dikembangkan oleh Dr.

Markus Hofmann dari Blanchardstown Institute of Technology dan Ralf Klinkenberg dari rapid-i.com memiliki tampilan GUI (graphical user interface) yang memudahkan pengguna dalam menggunakan perangkat lunak.

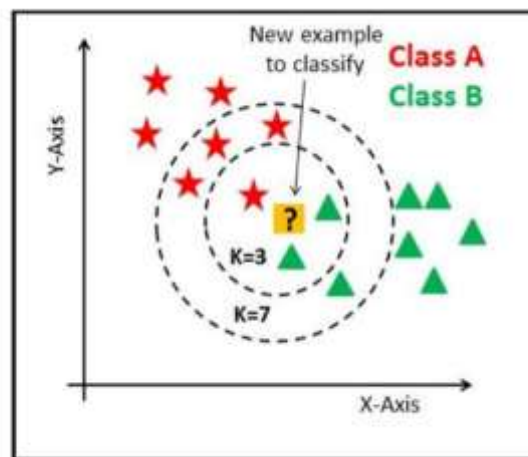
Perangkat lunak ini bersifat open source dan dibuat menggunakan program Java di bawah Lisensi Publik GNU. Rapid Miner dapat berjalan di sistem operasi apa pun. Dengan Rapid Miner, Anda tidak memerlukan pengetahuan pemrograman khusus karena semua fitur sudah disertakan.

Rapid Miner khusus digunakan untuk penambangan data. Model yang ditawarkan juga sangat banyak dan lengkap, seperti: Model Bayesian, pemodelan, induksi pohon, Nural Network, dan lain-lain.

2.2.10 Algoritma

Algoritma merupakan metode efektif yang dinyatakan sebagai deret hingga. Algoritma juga merupakan kumpulan perintah untuk memecahkan masalah. Perintah-perintah ini dapat diterjemahkan langkah demi langkah dari awal hingga akhir. Masalah dapat berupa masalah apa saja, asalkan untuk setiap masalah, kriteria kondisi awal harus dipenuhi sebelum algoritma dijalankan. Algoritma ini juga memiliki proses pengulangan (iterasi) dan pengambilan keputusan sampai keputusan tersebut selesai. Tidak peduli sistem komputer atau bahasa pemrograman apa yang Anda gunakan, algoritma adalah studi abstrak. Algoritma yang berbeda dapat diterapkan untuk masalah dengan kriteria yang sama. Kompleksitas suatu algoritma adalah ukuran seberapa banyak perhitungan yang digunakan algoritma untuk menyelesaikan masalah. Secara informal, algoritma yang dapat menyelesaikan masalah dalam

waktu yang relatif singkat memiliki kompleksitas yang rendah, sedangkan algoritma yang dapat menyelesaikan masalah dalam waktu yang lama memiliki kompleksitas yang tinggi. Kesulitan masalah ini terletak pada sulitnya memahami algoritma dan solusi dari masalah yang dihadapi, serta sulitnya membayangkan struktur data yang akan digunakan. Ketika memahami solusi suatu masalah, jika masalah tersebut dapat ditampilkan dalam bentuk visual dan gambar, kita akan lebih mudah mengingat dan memahaminya, yang membuat presentasi menjadi lebih menarik.



GAMBAR 2.4 Algoritma KNN

2.2.11 Machine Learning

Sebagai kombinasi kecerdasan buatan dan statistik, machine learning telah terbukti menjadi bidang penelitian yang bermanfaat, yang telah menghasilkan banyak masalah dan solusi yang berbeda untuk algoritma. Tujuan dari algoritma ini, kumpulan data pelatihan yang tersedia, dan strategi

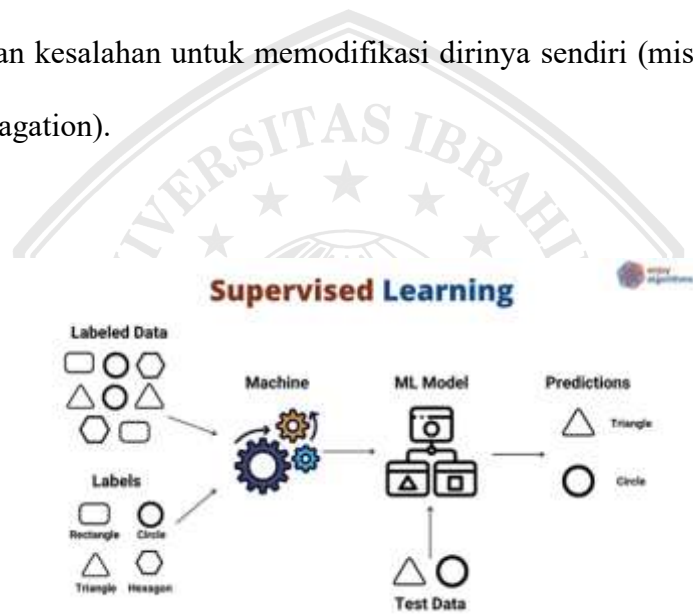
pembelajaran serta representasi data bervariasi. Namun, semua algoritma ini belajar dengan mencari ruang n-dimensi dari kumpulan data yang diberikan untuk menemukan generalisasi yang dapat diterima. Salah satu tugas machine learning paling dasar adalah pembelajaran mesin induktif, di mana generalisasi diperoleh dari sekumpulan sampel dan diformalkan menggunakan teknik dan model yang berbeda. Kita dapat mendefinisikan pembelajaran induktif sebagai proses menggunakan sejumlah pengamatan atau pengukuran input dan output sistem yang terbatas untuk memperkirakan ketergantungan input- output yang tidak diketahui atau struktur sistem. Dalam teori pembelajaran induktif, semua data dalam proses pembelajaran diatur, dan untuk setiap pasangan input - output, kami menggunakan istilah sederhana yang dikenal sebagai sampel.

Jenis-jenis machine learning diantaranya yaitu:

- Federated Learning
- Transfer Learning
- Deep Learning
- Self-Supervised Learning
- Reinforcement Learning
- Semi-Supervised Learning
- Unsupervised Learning
- Supervised Learning

Pada penelitian ini yaitu menggunakan model Supervised Learning (Pembelajaran terawasi) yang mana pengertian dari Supervised Learning adalah Model belajar dari sekumpulan data berlabel dan kemudian digunakan

untuk memprediksi kejadian di masa depan. Untuk proses pelatihan, inputnya adalah sekumpulan data pelatihan yang diketahui dengan label yang sesuai, dan algoritma pembelajaran menghasilkan fungsi inferensi, dan akhirnya memprediksi beberapa pengamatan baru yang tidak terlihat yang dapat ditetapkan ke model. Model dapat memberikan target untuk setiap input baru setelah dilatih sepenuhnya. Algoritma ini juga dapat membandingkan outputnya dengan output yang diinginkan (label kebenaran dasar) dan menemukan kesalahan untuk memodifikasi dirinya sendiri (misalnya melalui back-propagation).



Gambar 2.6 Supervised Learning

Model su

1. Regresi: Masalah regresi adalah ketika variabel
2. Klasifikasi: Masalah klasifikasi adalah ketika variabel keluaran adalah kategori, mis. "penyakit" / "tidak ada penyakit".

Output merupakan nilai lanjutan yang nyata, misalnya prediksi harga saham Beberapa contoh model yang termasuk dalam supervised ini adalah sebagai berikut: SVC, LDA, SVR, KNN, regression, random forests dll.

BAB III

METODELOGI PENELITIAN

3.1 Metodologi Penelitian

Metode penelitian adalah cara sistematis dalam mengumpulkan dan menganalisis data dengan tujuan mendapatkan informasi atau menjawab pertanyaan penelitian. Metode ini sangat penting untuk memastikan hasil penelitian valid dan dapat diandalkan. Ada berbagai pendekatan dalam metode penelitian, tergantung pada tujuan, jenis data, dan bidang studi. Pada penelitian ini peneliti menggunakan metode penelitian *kualitatif* yang bertujuan untuk Memahami fenomena dari sudut pandang peserta dan mendapatkan wawasan tentang bagaimana mereka mengartikan pengalaman mereka dan Menggali makna di balik perilaku, pengalaman, dan interaksi dalam konteks sosialnya serta Menyelidiki proses sosial dan perubahan yang terjadi dari waktu ke waktu. Berikut ini merupakan langkah langkah dalam menggunakan metode kualitatif:

1. Identifikasi Masalah Penelitian:

Tentukan topik atau fenomena yang ingin diteliti. Fokuskan pada aspek-aspek yang ingin dipahami lebih dalam.

2. Tinjauan Literatur:

Teliti literatur yang relevan untuk memahami konteks dan penelitian sebelumnya terkait dengan topik. Ini membantu dalam merumuskan pertanyaan penelitian dan membangun dasar teoretis.

3. Perumusan Pertanyaan Penelitian:

Buat pertanyaan penelitian yang terbuka dan eksploratif. Pertanyaan ini harus mampu menggali pengalaman, pandangan, atau persepsi subjek penelitian.

4. Desain Penelitian:

Pilih metode pengumpulan data yang sesuai, seperti wawancara mendalam, diskusi kelompok fokus (focus group discussion), atau observasi. Rancang instrumen pengumpulan data seperti panduan wawancara.

5. Sampling:

Tentukan sampel penelitian dengan pendekatan purposive atau snowball sampling, yang memilih peserta berdasarkan kriteria tertentu atau berdasarkan rekomendasi dari peserta awal.

6. Pengumpulan Data:

Lakukan pengumpulan data sesuai dengan metode yang dipilih. Pastikan prosesnya dilakukan secara sistematis dan etis.

7. Transkripsi Data:

Jika menggunakan wawancara atau diskusi, transkripkan data audio ke dalam bentuk teks. Transkripsi ini penting untuk analisis data lebih lanjut.

8. Analisis Data:

Lakukan analisis data dengan metode seperti coding, kategori, dan tema, Identifikasi pola, hubungan, dan tema yang muncul dari data.

9. Interpretasi Data:

Tafsirkan hasil analisis untuk memahami makna dan implikasi dari data yang dikumpulkan. Hubungkan temuan dengan teori dan literatur yang ada.

10. Penulisan Laporan:

Susun laporan penelitian yang mencakup latar belakang, metodologi, temuan, dan kesimpulan. Jelaskan bagaimana hasil penelitian dapat diterapkan atau dikembangkan lebih lanjut.

3.1 Lokasi Penelitian

Adapun untuk lokasi penelitian pada kasus ini yaitu di Madrasah Ibtida'iyah Al Fagiri Dusun Krajan RT.03 RW.04, desa Pace. Kec. Silo, kab. Jember

3.2 Bahan Penelitian

Pada penelitian ini peneliti menggunakan bahan untuk diolah menjadi data, yaitu data kehadiran guru dan data hasil pengajaran di Madrasah Ibtida'iyah Al Fagiri Pace Silo Jember.

3.3 Metode Pengumpulan Data

Adapun teknik pengumpulan data pada penelitian ini yaitu:

a. Observasi

Observasi merupakan teknik pengumpulan data dengan mengamati secara langsung aktivitas yang terjadi di lapangan agar mengetahui alur proses sistem yang berjalan yang bertujuan untuk memperoleh data atau informasi tentang objek penelitian.

b. Wawancara

Wawancara adalah teknik pengumpulan data dimana peneliti dan pemangku kepentingan berinteraksi secara langsung melalui pertanyaan-pertanyaan yang telah disepakati sebelumnya. Untuk mengembangkan sistem informasi yang sesuai dengan kebutuhan, peneliti melakukan wawancara dengan pihak-pihak terkait mengenai alur kerja dan aturan terkait sistem yang akan dibuat. Wawancara merupakan suatu proses komunikasi yang melibatkan dua pihak atau lebih yang bertanya dan menjawab pertanyaan untuk memperoleh informasi tertentu. Dalam konteks ini, biasanya ada pewawancara yang mengajukan pertanyaan dan satu atau lebih sumber memberikan jawaban.

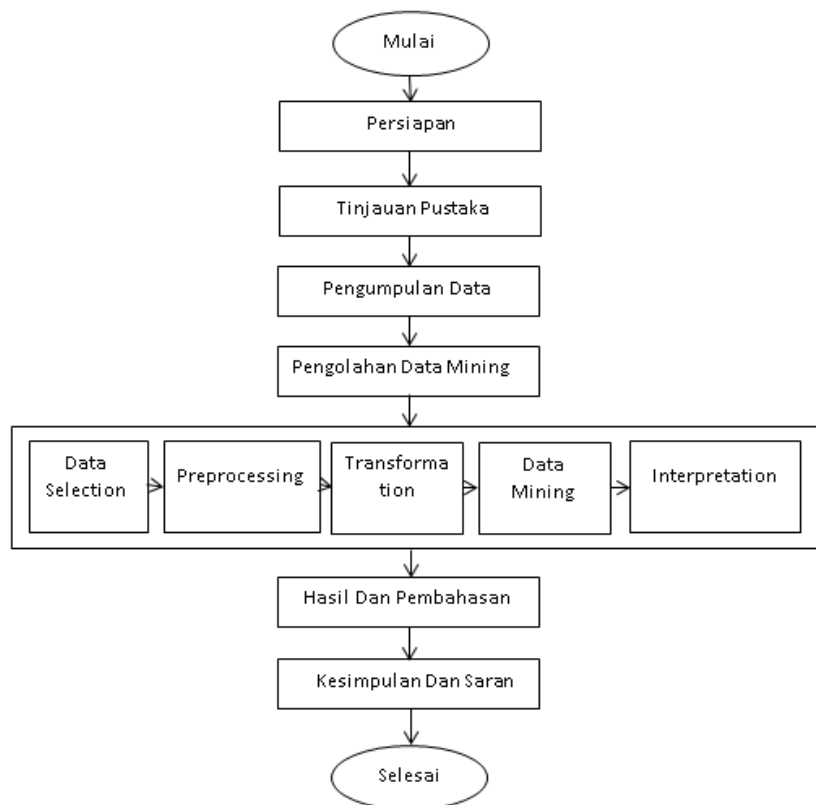
c. Data Sekunder

Data sekunder merupakan suatu bentuk data historis yang berisi informasi tentang variabel-variabel yang sebelumnya telah dikumpulkan dan diarsipkan oleh entitas lain. Data sekunder penelitian ini diperoleh dari sekolah itu sendiri dan penulis menggunakan data sekunder internal yang meliputi data yang dikumpulkan dari masing-masing sekolah seperti, Hasil kelas, laporan kehadiran, evaluasi keikutsertaan dalam kegiatan sekolah, dan laporan pengetahuan kelas..

3.4 Tahapan Penelitian

Adapun proses dan tahapan-tahapan dalam penelitian ini sebagai berikut:

- a. Persiapan: pada tahapan ini merupakan tahapan mengenai objek tempat penelitian dan subjek yang diteliti yaitu pada Madrasah Ibtida'iyah Al Fagiri dan Data laporan Kehadiran Guru
- b. Tinjauan Pustaka: pada tahapan ini dilakukan telaah dan literatur studi yang berkaitan dengan penilaian kinerja dan yang berhubungan
- c. Pengumpulan Data: tahapan ini merupakan proses memperoleh data melalui wawancara, observasi kepada bagian petugas
- d. Pengolahan Data Mining: data yang sudah diperoleh kemudian diolah sesuai dengan tahapan Knowledge Discovery in Database (KDD)
- e. Hasil dan Pembahasan: pada tahapan ini merupakan proses pengolahan data yang diperoleh dengan menggunakan metode K-Nearest Neighbor (K-NN)
- f. Kesimpulan dan Saran: semua rangkaian penelitian dari awal sampai terakhir hingga selesai akan menghasilkan kesimpulan yang akan dituangkan dibagian terakhir



Gambar 3.1 Tahapan Penelitian

3.5 Pengolahan Data Mining

Proses data mining pada penelitian ini mengikuti setiap fase Knowledge Discovery in Database (KDD) dan menghasilkan informasi pada tahap-tahap berikutnya dengan urutan yang telah ditentukan sebagai berikut:

3.5.1 Data Selection

Seleksi data adalah proses memilih subkumpulan data yang relevan dan berkualitas tinggi dari kumpulan data yang besar untuk dianalisis lebih lanjut.

Data yang dipilih digunakan untuk proses data mining dan disimpan dalam file terpisah dari database produksi. Data yang digunakan dalam penelitian ini berasal dari Madrasah Ibtidiah Al-Fagiri yaitu data laporan kehadiran dan data penilaian hasil pembelajaran.

Seluruh atribut yang dimasukkan dalam data evaluasi kinerja ini meliputi nama guru, tingkat ketidakhadiran, peringkat pembelajaran, dan bulan.

Tiga field yang dipilih dari seluruh data yang ada dan digunakan untuk proses Knowledge Discovery in Database (KDD) diantaranya adalah sebagai berikut:

1. Nama Guru merupakan atribut yang terdapat pada tabel data penilaian kinerja yang berisi informasi tentang nama guru yang ada di Madrasah Ibtida'iyah Al Fagiri.
2. Absensi adalah atribut yang terdapat dalam tabel penilaian kinerja yang berisi informasi seberapa banyaknya kehadiran guru.
3. Bulan adalah atribut yang berada pada tabel data penilaian yang berisikan informasi bulan laporan.

3.5.2 Preprocessing

Preprocessing atau pra-pemrosesan data adalah proses analisis data dan data mining yang bertujuan untuk mempersiapkan data mentah menjadi bentuk yang lebih bersih dan terstruktur. selanjutnya dilakukan data cleaning untuk menghasilkan dataset yang bersih sehingga dapat digunakan dalam tahap berikutnya yaitu mining Proses ini sangat penting untuk memastikan bahwa

data yang digunakan dalam analisis atau model machine learning berkualitas tinggi dan dapat memberikan hasil yang akurat. Berikut merupakan penjelasan dari 2 hal tersebut:

1. Integrasi Data adalah proses menggabungkan data dari berbagai sumber menjadi satu dataset yang kohesif dan konsisten. Tujuannya adalah untuk mengintegrasikan informasi yang tersebar di berbagai sistem, database, atau format sehingga dapat digunakan secara efektif untuk analisis, pelaporan, atau aplikasi lain.
2. Pembersihan Data (Data Cleaning) adalah proses manajemen data yang bertujuan untuk memperbaiki kualitas data dengan menghapus atau memperbaiki data yang tidak akurat, tidak konsisten, atau tidak lengkap. Pembersihan data memastikan bahwa data yang digunakan dalam analisis, pelaporan, atau model machine learning adalah akurat, konsisten, dan dapat diandalkan.

3.5.3 Transformation

Transformation dalam konteks pengolahan data adalah proses mengubah data dari bentuk atau format yang satu ke bentuk atau format yang lain untuk memudahkan analisis, integrasi, atau pemodelan. Transformation sering kali dilakukan setelah tahap pembersihan data dan sebelum data digunakan untuk analisis lebih lanjut atau dimuat ke dalam sistem data warehouse. Tujuan utama dari transformation adalah untuk memastikan bahwa data

memiliki struktur yang sesuai, konsisten, dan berguna untuk kebutuhan analisis.

3.5.4 Data Mining

Data Mining adalah proses menemukan pola, hubungan, atau informasi berguna dari kumpulan data besar menggunakan teknik analisis statistik, algoritma machine learning, dan alat analisis data. Tujuan dari data mining adalah untuk mengeksplorasi data dan mengidentifikasi informasi tersembunyi yang dapat membantu dalam pengambilan keputusan, prediksi, dan pengembangan strategi bisnis. Metode yang digunakan dalam penelitian ini adalah metode K-Nearest Neighbor (KNN). Atribut untuk metode ini diinisialisasi sebagai k. Artinya, banyaknya tetangga yang dijadikan acuan dalam KNN adalah nilai k yang merupakan bilangan bulat positif kecil dan ganjil. Langkah-langkah kerja metode ANN pada penelitian ini adalah sebagai berikut:

- a. Tentukan nilai k. Meskipun belum ada aturan baku dalam menentukan nilai k yang digunakan untuk klasifikasi, namun dalam penelitian ini digunakan nilai k 3, 5, 7, 9, 11, 13, dan 15.
- b. Menghitung jarak antara data latih dan data uji. Metode perhitungan jarak yang digunakan pada metode KNN adalah Euclidean distance. dengan rumus sebagai berikut:

$$\sqrt{\sum_{i=1}^K (X_i - Y_i)^2}$$

Contoh penghitungan jarak pada KNN ini menggunakan jarak Euclidean dijelaskan pada bab berikutnya dengan menggunakan data nyata.

- c. Mengurutkan data perhitungan. Jarak yang ditentukan kemudian diurutkan dari yang terdekat hingga terjauh (dalam urutan menaik).
- d. Menentukan dataset hasil pengujian berdasarkan label mayoritas tetangga terdekat..

3.5.5 Interpretation

Pola informasi yang dihasilkan dari proses data mining perlu disajikan dalam format yang mudah dipahami oleh pemangku kepentingan. Fase ini melibatkan pemeriksaan untuk melihat pola dan informasi yang ditemukan konsisten dengan fakta dan hipotesis yang ada sebelumnya. Pada fase ini, pola penilaian kinerja ditemukan dari proses data mining menggunakan K-Nearest Neighbors. Informasi yang diperoleh dari proses data mining berupa aturan-aturan yang diperoleh dari proses perhitungan K-NN.

BAB IV

HASIL DAN PEMBAHASAN

4.1 Gambaran Umum Lokasi Penelitian

4.1.1 Sejarah Singkat Instansi

Madrasah Ibtida'iyah Al fagiri berlokasi di dusun Krajan desa Pace kecamatan Silo kabupaten Jember. Madrasah ini didirikan pada tahun 2002 yang telah disebutkan secara spesifik dalam dokumen resmi. Namun sebagai salah satu lembaga pendidikan dasar Islam, sekolah ini mempunyai sejarah yang berakar pada perkembangan pendidikan agama Islam di daerah. Sekolah ini didirikan oleh tokoh masyarakat setempat dan para pendidik dengan tujuan menyelenggarakan pendidikan dasar berdasarkan kurikulum agama Islam dengan tetap mengintegrasikan nilai-nilai pendidikan umum. Madrasah Al fagiri menanamkan ajaran Islam sejak usia dini, memberikan siswa landasan akademik dan spiritual yang kuat dan mempersiapkan mereka untuk pendidikan tinggi. Seiring berjalannya waktu, Madrasah Ibtida'iyah Al fagiri telah mengalami berbagai perubahan dan pengembangan, termasuk dari segi kurikulum dan fasilitas. Madrasah Ibtida'iyah Al fagiri berupaya untuk memberikan pendidikan berkualitas tinggi yang menggabungkan kurikulum umum dan agama sehingga siswa tidak hanya memperoleh pengetahuan akademis tetapi juga memahami ajaran Islam secara mendalam. Tidak jauh berbeda dengan sekolah-sekolah lain pada umumnya, Madrasah Ibtida'iyah Al fagiri juga memiliki fasilitas standar seperti ruang kelas, perpustakaan, dan taman bermain. Dalam beberapa tahun terakhir, upaya telah dilakukan untuk meningkatkan fasilitas dan sumber daya untuk mendukung proses belajar mengajar yang lebih efektif. Madrasah

Ibtida'iyah Al fagiri Krajan Pace Silo tetap berkomitmen meningkatkan mutu pendidikan dan memberikan kontribusi positif terhadap pembangunan pendidikan di Kabupaten Jember khususnya kecamatan Silo. Diharapkan dengan dukungan masyarakat setempat dan pihak sekolah, madrasah ini mampu mengembangkan dan melayani kebutuhan pendidikan anak-anak di daerah tersebut.

4.1.2 Visi

Mewujudkan generasi yang mandiri, unggul dalam prestasi dan berakhlakul karimah.

4.1.3 Misi

- a. Menyelenggarakan Pendidikan akademis, religious, mandiri dan terampil.
- b. Menyelenggarakan pembelajaran yang efektif dan kondusif dalam pendidikandan pengembangan sumber daya manusia.
- c. Mengadakan pembinaan agama secara berkelanjutan, bakat siswa melalui wadah ekstra kulikuler (Pramuka, Olahraga, TPQ, Tahfidz) sehingga menghasilkan siswa yang beriman, bertaqwa, dan berakhlakul karimah, menguasai pengetahuan luas dan seimbang serta terampil membaca Al qur'an dan memahaminya.

4.1.4 Tujuan

1. Pendidikan Agama yang Kuat: memberikan pendidikan Islam yang menyeluruh dan komprehensif yang mencakup pemahaman tentang Al-

Qur'an, hadis, akhlak, dan fiqhān, sehingga siswa dapat membangun landasan keimanan dan ketaqwaan yang kokoh.

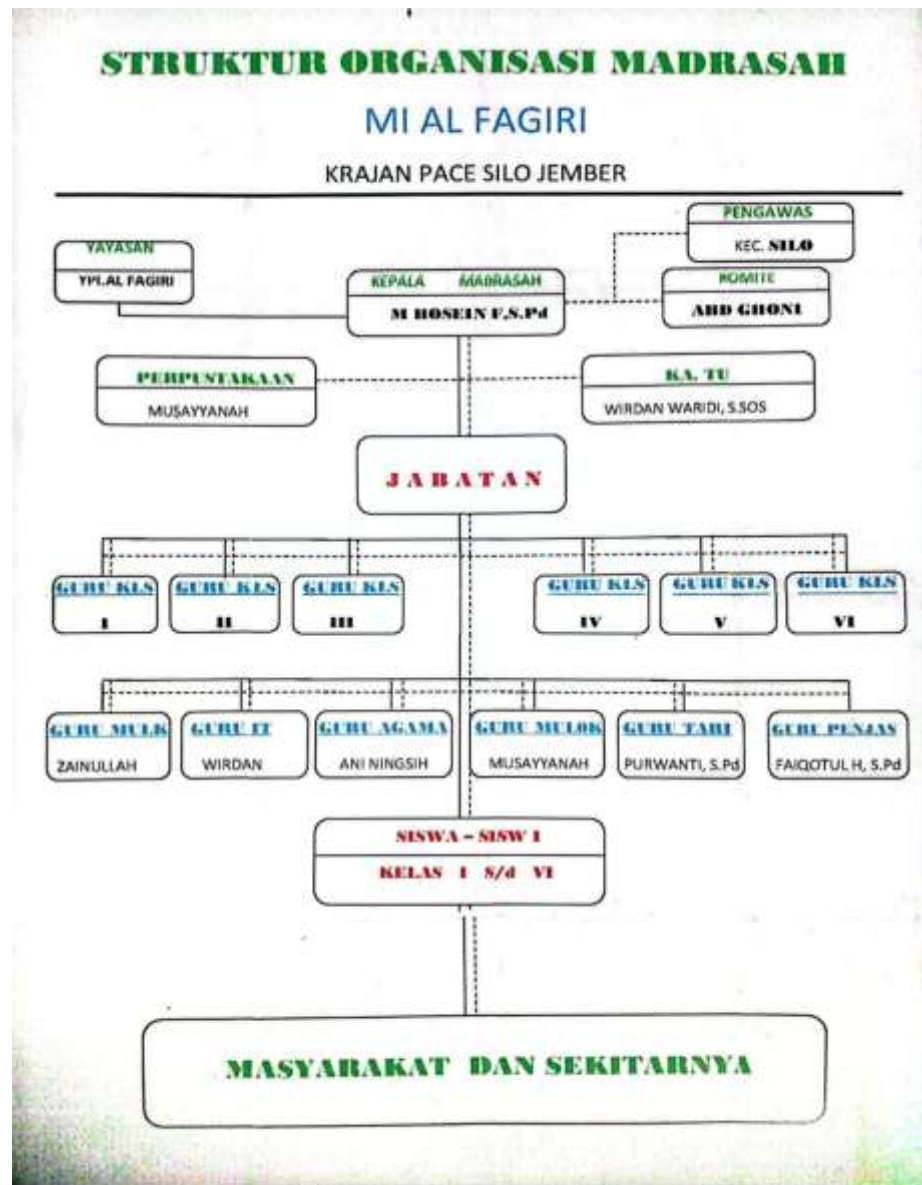
2. Pengembangan Akademik: Memberikan pelatihan akademik berkualitas tinggi dalam mata pelajaran umum seperti Matematika, Bahasa Indonesia, Bahasa Inggris, dan Ilmu Pengetahuan Sosial untuk mempersiapkan siswa untuk studi akademik di tingkat lanjutan.
3. Pendidikan Karakter : Mendidik peserta didik tentang pengembangan karakter dan akhlak mulia melalui pembelajaran dan praktik sehari-hari agar menjadi manusia yang berakhlak mulia, berintegritas, dan memiliki rasa tanggung jawab.
4. Keterampilan Sosial dan Kepemimpinan: Membantu siswa mengembangkan keterampilan sosial dan kepemimpinan melalui kegiatan ekstrakurikuler, kerja kelompok, dan partisipasi dalam berbagai program sekolah.
5. Kesehatan dan Kesejahteraan: Menyediakan lingkungan yang aman dan sehat yang mempertimbangkan aspek kesehatan fisik dan mental siswa serta mengedepankan gaya hidup sehat.
6. Pengembangan Kreativitas: Mendorong siswa untuk mengekspresikan kreativitasnya melalui berbagai kegiatan seni, budaya, dan inovasi, memberikan kesempatan untuk mengembangkan bakat dan minat individu.
7. Integrasi ke dalam masyarakat: Membangun hubungan yang harmonis dengan masyarakat sekitar dan melibatkan peserta didik dalam proses

pendidikan agar mereka memahami perannya dalam masyarakat dan dapat berkontribusi secara aktif.

Dengan cita-cita tersebut, MI Al-Fagiri bertujuan untuk melahirkan generasi yang tidak hanya cerdas secara akademis tetapi juga dibekali landasan agama yang kuat dan akhlak yang baik, siap menghadapi berbagai tantangan masa depan.



4.1.5 Struktur



Gambar 4.1. Struktur Madrasah Ibtida'iyah Al fagiri

4.2 Hasil

Setelah menganalisis terhadap desain dengan menggunakan tahapan data mining, maka dibuatlah peringkat penilaian kinerja guru terbaik pada Madrasah ibtida'iyah Al fagiri dengan menggunakan algoritma K-Nearest Neighbor. Untuk

mengetahui hasil prediksi penilaian kinerja guru terbaik berdasarkan data lima tahun terakhir dari tahun 2019-2023 dengan menggunakan aplikasi atau tools Rapidminer sebagai akses untuk mempermudah proses prediksi data mining yang akan menghasilkan informasi penilaian kinerja guru terbaik. Kemudian setelah melakukan proses prediksi dengan menggunakan data mining nilai akurasi yang di dapatkan yaitu sebesar 92,86%

4.3 Pembahasan

4.3.1 Data Selection

Pada penelitian ini data yang di gunakan adalah data persentase kehadiran yang diperoleh dari Madrasah Ibtidaiyah Al-Fagiri. Kemudian data itu diseleksi dan untuk diolah agar dapat mengetahui nilai dari kinerja para guru. Adapun atribut yang digunakan dalam penentuan penilaian kinerja guru terbaik adalah Nama, Kuantitas dan tahun Dapat dilihat pada Tabel 4.1.

NO	NAMA	TAHUN AJARAN	JUMLAH KEHADIRAN
1	M.HOSEN F, S.P.d	2020	354
2	M. ZAINULLAH MU	2020	350
3	KASIR MADLI	2020	278
4	ANI NINGSIH	2020	321
5	MUSAYYANAH	2020	298
6	WAHYUDI	2020	187
7	HUSNATUL PURWANTI	2020	321
8	NUR IMAMAH	2020	340
9	WIRDAN WARIDI	2020	276
10	NATAZA RHESA PURUHATA	2020	240
11	FIRDATUL HASANAH	2020	300

Tabel 4.1 Daftar hasir guru tahun ajaran 2019-2020

NO	NAMA	TAHUN AJARAN	JUMLAH KEHADIRAN
1	M.HOSEN F, S.P.d	2021	200
2	M. ZAINULLAH MU	2021	300
3	KASIR MADLI	2021	360
4	ANI NINGSIH	2021	360
5	MUSAYYANAH	2021	300
6	WAHYUDI	2021	200
7	HUSNATUL PURWANTI	2021	250
8	NUR IMAMAH	2021	360
9	WIRDAN WARIDI	2021	150
10	NATAZA RHESA PURUHATA	2021	200
11	FIRDATUL HASANAH	2021	300

Tabel 4.2 Daftar hasir guru tahun ajaran 2020-2021

NO	NAMA	TAHUN AJARAN	JUMLAH KEHADIRAN
1	M.HOSEN F, S.P.d	2022	160
2	M. ZAINULLAH MU	2022	360
3	KASIR MADLI	2022	250
4	ANI NINGSIH	2022	300
5	MUSAYYANAH	2022	200
6	WAHYUDI	2022	350
7	HUSNATUL PURWANTI	2022	150
8	NUR IMAMAH	2022	350
9	WIRDAN WARIDI	2022	300
10	NATAZA RHESA PURUHATA	2022	250
11	FIRDATUL HASANAH	2022	200

Tabel 4.3 Daftar hasir guru tahun ajaran 2021-2022

NO	NAMA	TAHUN AJARAN	JUMLAH KEHADIRAN
1	M.HOSEN F, S.P.d	2023	360
2	M. ZAINULLAH MU	2023	350

3	KASIR MADLI	2023	300
4	ANI NINGSIH	2023	320
5	MUSAYYANAH	2023	280
6	WAHYUDI	2023	300
7	HUSNATUL PURWANTI	2023	350
8	NUR IMAMAH	2023	300
9	WIRDAN WARIDI	2023	250
10	NATAZA RHESA PURUHATA	2023	350
11	FIRDATUL HASANAH	2023	250

Tabel 4.4 Daftar hasir guru tahun ajaran 2022-2023

NO	NAMA	TAHUN AJARAN	JUMLAH KEHADIRAN
1	M.HOSEN F, S.P.d	2024	256
2	M. ZAINULLAH MU	2024	300
3	KASIR MADLI	2024	344
4	ANI NINGSIH	2024	320
5	MUSAYYANAH	2024	280
6	WAHYUDI	2024	300
7	HUSNATUL PURWANTI	2024	350
8	NUR IMAMAH	2024	300
9	WIRDAN WARIDI	2024	250
10	NATAZA RHESA PURUHATA	2024	350
11	FIRDATUL HASANAH	2024	353

Tabel 4.5 Daftar hasir guru tahun ajaran 2023-2024

4.3.2 Preprocessing

Tahap preprocessing dilakukan untuk pengelompokan jenis jumlah kehadiran berdasarkan jumlah kehadiran tiap bulan dan tahun untuk mempermudah dalam proses perhitungan prediksi. Setelah data terkelompokan

lalu semua dijumlahkan sehingga menjadi data penilaian kinerja semua guru.

Dapat dilihat pada beberapa tabel dibawah ini :

NO	BULAN	THN-2019	THN-2020	THN-2021	THN-2022	THN-2023
1	JANUARI	21	27	25	30	25
2	FEBRUARI	19	18	30	25	28
3	MARET	16	15	25	22	17
4	APRIL	16	29	20	23	19
5	MEI	22	30	25	24	27
6	JUNI	24	18	25	22	20
7	JULI	16	15	24	22	25
8	AGUSTUS	17	19	25	26	30
9	SEPTEMBER	15	25	30	20	24
10	OKTOBER	15	15	29	25	20
11	NOVEMBER	30	20	26	28	24
12	DESEMBER	27	25	28	23	19

Tabel 4.2 Sampel Data kehadiran guru

4.3.3 Transformation

Pada tahapan transformasi ini, hasil dari pengelompokan data yang diolah digunakan sebagai data latih. Saat membuat data pelatihan berdasarkan data yang ada, Anda harus memilih data terlebih dahulu untuk ditentukan atribut mana yang paling berdampak terhadap kinerja guru. Hal ini disebut data target. Data target adalah data yang memuat atribut-atribut yang relevan dan mendukung proses data mining.

. Data pelatihan yang digunakan ditunjukkan pada tabel. 4.6.

No	Bulan-1	Bulan-2	Bulan-3	Bulan-4	Bulan-5
1	21	27	25	30	25
2	19	18	30	25	28
3	16	15	25	22	17
4	16	29	20	23	19
5	22	30	25	24	27
6	24	18	25	22	20
7	16	15	24	22	25
8	17	19	25	26	30
9	15	25	30	20	24
10	15	15	29	25	20
11	30	20	26	28	24
12	27	25	28	23	19

Tabel 4.7 Data Training Kehadiran Guru

Berdasarkan Tabel 4.7, data pelatihan yang digunakan dalam penelitian ini hanya data kehadiran guru selama lima tahun terakhir yaitu dari tahun 2019 hingga tahun 2023 yang merupakan data pelatihan. Data pelatihan dikelompokkan menjadi lima kolom. Data masukannya adalah data kehadiran guru dari bulan pertama sampai dengan bulan ke 5. Kemudian dijadikan data masukan, misalnya bulan ke 6, sebagai data target atau keluaran hingga batas data yang ada tercapai. Data training yang digunakan untuk prediksi ditunjukkan pada tabel berikut.

BULAN-1	BULAN-2	BULAN-3	BULAN-4	BULAN-5
27	28	26	30	17
21	18	24	24	18
24	30	29	28	29
22	21	20	20	24
26	25	22	29	18
16	29	25	26	23
16	25	21	29	16
28	21	29	22	16

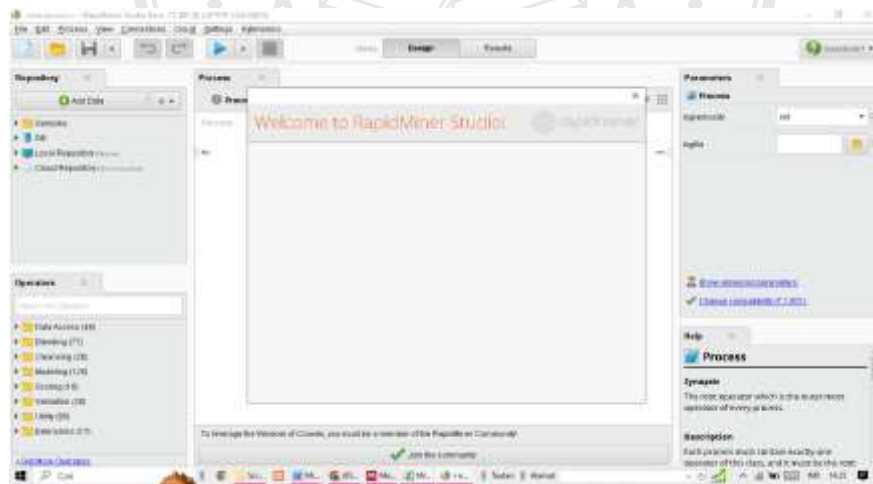
23	27	25	22	17
27	15	28	26	21
21	18	28	29	29
25	23	24	22	26

Tabel 4.8 Data testing kehadiran guru

Tabel 4.8 menunjukkan bahwa data yang digunakan untuk pengujian adalah data Penilaian berdasarkan kehadiran bulan Januari 2019 hingga Mei 2019 yang memberikan prediksi peringkat kinerja guru pada bulan berikutnya.

4.3.4 Implementasi Rapidminer

RapidMiner adalah sebuah software data mining pengolahan data set untuk analisis data, pembelajaran mesin, dan analitik prediktif. Platform ini memungkinkan pengguna untuk memproses data, membuat model analitik, dan menerapkan hasil analisis dengan antarmuka grafis yang intuitif.



Gambar 4.1 Tampilan awal Rapidminer studio

Pada gambar 4.1 merupakan tampilan awal dari rapidminer studio dan tools yang digunakan pada rapidminer ini yaitu tipe 7.1.001

Pada gambar 4.3 merupakan proses import data excel kedalam rapidminer dengan cara draghere read excel lalu pilih data yang akan kita masukkan



Gambar 4.4 Penentuan Label

Pada gambar 4.4 merupakan proses setelah memasukkan data kedalam rapidminer sehingga muncul keterangan atribut seperti gambar diatas yang mana dirubah menjadi label/keterangan dengan tujuan agar prediksi penilaian kinerja guru ketemu setelah di run/dijalankan nantinya



Gambar 4.5 Proses K-NN

Selanjutnya yaitu pemrosesan data menggunakan metode K-NN dan karena ini merupakan prediksi maka metode yang paling tepat yaitu menggunakan metode K-NN



Gambar 4.6 Pengulangan import data

Karena metode k-nn merupakan pengujian klasifikasi data, maka data harus ada 2 data yang diimport tujuannya agar ketemu antara data training dan data uji seperti gambar 4.6 diatas

BAB V

PENUTUP

5.1 Kesimpulan

Setelah melakukan beberapa tahap, peneliti menggunakan pemodelan algoritma K-NN menggunakan pengolahan data berdasarkan tahap Knowledge Discovery in Database (KDD). Berdasarkan metode klasifikasi dan hasil perhitungan data mining menggunakan algoritma K-NN diperoleh nilai akurasi prediksi penilain kinerja guru sebesar 92,86%.

5.2 Saran

Berdasarkan evaluasi terhadap proses dan hasil dari imolementasi perhitungan menggunakan metode K-NN pada penilaian kinerja guru ini, maka dapat diberikan beberapa saran yang dapat digunakan untuk pengembangan lebih lanjut dari pengimplementasian metode K-NN tersebut, agar hasil yang diperoleh lebih maksimal.

1. Penelitian lebih lanjut dapat mengembangkan suatu sistem untuk membantu menghitung penilaian kinerja guru. Agar instansi terkait lebih mudah menentukan penilaian dan evaluasi kinerja guru.
2. Untuk penelitian selanjutnya, kita juga dapat Menggabungkan dua teknik algoritma data mining dan menambahkan teknik data mining agar dapat menghasilkan yang beragam dan nilai informasi dan akurasi yang jauh lebih bagus.

DAFTAR PUSTAKA

- [1] S. P. M. P. Prof. Dr. H. Elfrianto Nst, M. P. Dr. Indra Utama, Z. Amin M, and S. P. M. P. Dr. Muhammad Arifin, *Manajemen Kinerja Guru Dalam Konteks Kurikulum Merdeka; Peningkatan Efektivitas Pembelajaran*. umsu press, 2024. [Online].
- [2] P. K. Neighbor, “Algoritma K-Nearest Neighbors (KNN) – Pengertian dan Penerapan Pengertian K-Nearest Neighbor (KNN),” 2023.
- [3] M. S. Prof. Dr. H. Elfrianto Nst, S.Pd., M.Pd, Dr. Indra Utama, M.Pd, Akmaluddin, Dr. Zulkifli Amin, *Manajemen Kinerja Guru Dalam Konteks Kurikulum Merdeka; Peningkatan Efektivitas Pembelajaran*. Medan: umsu press, 2024. [Online].
- [4] S. Wiyono, P. Harapan Bersama, and T. Abidin, “Implementasi Algoritma K-Nearest Neighbor (KNN) Ke Prediksi Kinerja Siswa,” *J. SIMETRIS*, vol. 9, no. 2, pp. 873–878, 2018, [Online].
- [5] I. Made *et al.*, *Implementasi Algoritma KNN untuk Memprediksi Performa Siswa Sekolah*, vol. 1, no. 3. 2023. [Online].
- [6] R. C. S. Hariyono *et al.*, “Buku Ajar Pengantar Basis Data.” p. 165, 2023.
- [7] A. Azis, A. T. Zy, and A. S. Sunge, “Prediksi Penjualan Obat Dan Alat Kesehatan Terlaris Menggunakan Algoritma K-Nearest Neighbor,” *J. Teknol. Dan Sist. Inf. Bisnis*, vol. 6, no. 1, pp. 117–124, 2024, doi:

10.47233/jteksis.v6i1.1078.

- [8] A. W. Zunan Setiawan, Muhammad Fajar , Arif Mudi Priyatno, Anggi Yhurinda Perdana Putri, Mediana Aryuni , Siti Yuliyanti, Harya Widiputra , Budanis Dwi Meilani, Rohmat Nur Ibrahim, Rezania Agramanisti Azdy, Satrio Junaidi, “Buku Ajar Data Mining.” p. 197, 2023.
- [9] M. A. Fitriani Dwi Ramadhani, *SISTEM PREDIKSI PENJUALAN DENGAN METODE SINGLE EXPONENTIAL SMOOTHING DAN TREND PARABOLIK*. Pascal Books, 2022. [Online].
- [10] “Cara_Mudah_Memahami_Algoritma_K_Nearest.”
- [11] M. R. N. , T. R. , Dwi Marisa Midyanti, “Penerapan Data Mining Menggunakan Metode K-Nearest Neighbor Untuk Klasifikasi Indeks Cuaca Kebakaran Berdasarkan Data Aws (Automatic Weather Station) (Studi Kasus: Kabupaten Kubu Raya),” *Coding J. Komput. dan Apl.*, vol. 6, no. 2, 2018, doi: 10.26418/coding.v6i2.26672.
- [12] R. Aprilian, R. Habibi, M. Y. H. Setyawan, and R. M. Awangga, *Algoritma KNN dalam memprediksi cuaca untuk menentukan tanaman yang cocok sesuai musim*. Kreatif, 2020. [Online].

CURRICULUM VITAE



Nama Lengkap : MUHAMMAD JUNDANUDDIN

NPM : 2020502025

Tempat, Tanggal Lahir : Jember, 12 September 2002

Program Studi : Sistem Informasi

Nama Orang Tua

Ayah : Zainulla MU

Ibu : Ani Ningsih

Latar Belakang Pendidikan

SD/MI : MI AL FAGIRI

SLTP/MTs : MTs "UNGGULAN" NURIS JEMBER

SLTA/MA : SMA IBRAHIMY SUKOREJO

Latar Organisasi : IKSASS

Alamat Rumah : Krajan Pace Silo Jember

No. Telepon : 085732135455

E-mail : jundannuddin@gmail.com

DOKUMENTASI

(Peneliti Bersama Kepala Sekolah Madrasah Ibtidaiyah Alfagiri Kecamatan Silo Kabupaten Jember)





(Ruang Guru Madrasah Ibtida'iyah Alfaqiri Kecamatan Silo Kabupaten Jember)



(Kantor Madrasah Ibtida'iyah Alfaqiri Kecamatan Silo Kabupaten Jember)