

**PENINGKATAN AKURASI KNN DENGAN MENGGUNAKAN
METODE *PRINCIPAL COMPONENT ANALYSIS*
PADA DATA JENIS JAMUR**

SKRIPSI

OLEH:

HARYA SYAFNELDI

188160105



**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS MEDAN AREA
MEDAN
2025**

UNIVERSITAS MEDAN AREA

© Hak Cipta Di Lindungi Undang-Undang

1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area

Document Accepted 26/8/25

Access From (repository.uma.ac.id)26/8/25

PENINGKATAN AKURASI KNN DENGAN MENGGUNAKAN
METODE *PRINCIPAL COMPONENT ANALYSIS*
PADA DATA JENIS JAMUR

SKRIPSI

Diajukan sebagai salah satu syarat untuk memperoleh
Gelar Sarjana di Fakultas Teknik
Universitas Medan Area

OLEH :

HARYA SYAFNELDI

188160105

PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS MEDAN AREA
MEDAN

2025

UNIVERSITAS MEDAN AREA

© Hak Cipta Di Lindungi Undang-Undang

1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area

Document Accepted 26/8/25

Access From (repository.uma.ac.id)26/8/25

HALAMAN PENGESAHAN

Judul Skripsi : *Peningkatan Akurasi KNN Dengan Menggunakan Metode
Principal Component Analysis Pada Data Jenis Jamur*

Nama : HARYA SYAFNELDI

NPM : 185160105

Fakultas : Teknik

Disetujui Oleh
Komisi Pembimbing



Nurul Khairina S. Kom, M. Kom

Pembimbing I



Dr. Riza M. M. S. Kom, M. Kom

Fakultas Teknik



Riza M. M. S. Kom, M. Kom

Ketua Prodi

Tanggal Lulus : 12 Maret 2025

HALAMAN PERNYATAAN

Saya menyatakan bahwa skripsi yang saya susun, sebagai syarat memperoleh gelar sarjana merupakan hasil karya tulis saya sendiri. Adapun bagian-bagian tertentu dalam penulisan skripsi ini yang saya kutip dari hasil karya orang lain telah dituliskan sumbernya secara jelas sesuai dengan norma, kaidah, dan etika penulisan ilmiah.

Saya bersedia menerima sanksi pencabutan gelar akademik yang saya peroleh dan sanksi-sanksi lainnya dengan peraturan yang berlaku, apabila di kemudian hari ditemukan adanya plagiat dalam skripsi ini.

Medan, 12 Maret 2025

Penulis,



HARYA SYAFNELDI

NPM : 188160105

HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR/SKRIPSI/TESIS UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademik Universitas Medan Area, saya yang bertanda tangan di bawah ini :

Nama : HARYA SYAFNELDI

NPM : 188160105

Program Studi : Informatika

Fakultas : Teknik

Jenis karya : Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Medan Area **Hak Bebas Royalti Noneksklusif**

(Non-exclusive Royalty-Free Right) atas karya ilmiah saya yang berjudul :

Peningkatan Akurasi KNN Dengan Menggunakan Metode *Principal Component Analysis* Pada Data Jenis Jamur

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Noneksklusif ini Universitas Medan Area berhak menyimpan, mengalihmedia/format-kan, mengelola dalam bentuk pangkalan data (database), merawat, dan memublikasikan tugas akhir/skripsi/tesis saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta. Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Kota Medan

Pada tanggal : 12 Maret 2025

Yang menyatakan



(HARYA SYAFNELDI)

RIWAYAT HIDUP

Penulis, Harya Syafneli, lahir di Pauh Kamar, Kecamatan Nan Sabaris, Kabupaten Padang Pariaman, Provinsi Sumatera Barat, pada tanggal 30 April 1999 pindah ke Kota Medan pada tahun 2000.

Penulis adalah putra pertama dari orang tua bernama Syafei dan Sinel dan memiliki empat saudara kandung, menempuh pendidikan di Kota Medan mulai dari pada Tahun 2005 SD di Yayasan Perguruan Pahlawan Nasional Medan dan tamat pada Tahun 2011, Tahun 2011 SMP di Yayasan Perguruan Pahlawan Nasional Medan dan tamat pada Tahun 2014, Tahun 2014 SMK di Tritech Informatika IT Medan dan tamat pada Tahun 2017 tamat dari SMK di Tahun 2018 Mendaftar sebagai Mahasiswa Fakultas Tekni Informatika di Universitas Medan Area tamat pada Tahun 2025.



ABSTRAK

Penelitian ini bertujuan untuk meningkatkan akurasi klasifikasi jamur menggunakan algoritma *K-Nearest Neighbor* (KNN) dengan bantuan *Principal Component Analysis* (PCA). KNN, meskipun sederhana, sering menghadapi tantangan dalam menangani dataset dengan dimensi tinggi yang mungkin mengandung fitur-fitur tidak relevan. PCA berperan sebagai teknik reduksi dimensi yang efektif untuk mengatasi masalah ini dengan mengidentifikasi dan mempertahankan fitur-fitur paling informatif. Hasil penelitian menunjukkan bahwa integrasi PCA dengan KNN berhasil meningkatkan akurasi klasifikasi jamur secara signifikan. Model KNN standar mencapai akurasi 94.96%, sementara KNN dengan PCA mencapai akurasi 97.45%. Peningkatan ini menegaskan bahwa PCA mampu menyaring fitur-fitur penting, memungkinkan KNN untuk bekerja lebih efisien. Penelitian ini menyimpulkan bahwa kombinasi PCA dan KNN merupakan pendekatan yang menjanjikan dalam meningkatkan akurasi klasifikasi, terutama pada dataset dengan dimensi tinggi.

Kata Kunci: *K-Nearest Neighbor*, *Principal Component Analysis*, Klasifikasi Jamur, Reduksi Dimensi

ABSTRACT

This research aimed to improve the accuracy of mushroom classification using the K-Nearest Neighbor (KNN) algorithm with the help of Principal Component Analysis (PCA). Although simple, KNN often faced challenges in handling highdimensional datasets that may contain irrelevant features. PCA served as an effective dimensionality reduction technique to overcome this issue by identifying and retaining the most informative features. The research results showed that integrating PCA with KNN significantly improved mushroom classification accuracy. The standard KNN model achieved 94.96% accuracy, while KNN with PCA reached 97.45%. This improvement confirmed that PCA successfully filtered important features, enabling KNN to work more efficiently. This research concluded that combining PCA and KNN was a promising approach to improving classification accuracy, especially for high-dimensional datasets.

Keywords: K-Nearest Neighbor, Principal Component Analysis, Mushroom Classification, Dimensionality Reduction

KATA PENGANTAR

Puji dan syukur penulis ucapkan atas kehadiran Allah SWT atas segala limpahan rahmat dan karunia yang diberikan sehingga penulis bisa menyelesaikan skripsi ini. Penulis menerima banyak bantuan, dukungan dan semangat selama menempuh pendidikan sampai dengan penyusunan skripsi ini. Oleh karena itu penulis menyampaikan ucapan terima kasih kepada :

1. Bapak Prof. Dr. Dadan Ramdan, M.Eng, M.Sc. selaku Rektor Universitas Medan Area Medan.
2. Bapak Dr. Eng. Supriatno, ST, MT selaku Dekan Fakultas Teknik Universitas Medan Area.
3. Bapak Rizki Muliono, S.Kom, M.Kom selaku Ketua Program Studi Teknik Informatika.
4. Ibu Nurul Khairina, S.Kom, M.Kom selaku Dosen Pembimbing yang telah memberikan bimbingan kepada penulis dalam penyelesaian skripsi.
5. Kedua orang tua tercinta, atas limpahan kasih sayang, dukungan dan doa yang telah diberikan kepada penulis.

Penulis juga menyadari bahwa penyusunan Tugas Akhir ini belum sempurna baik dalam penulisan maupun isi disebabkan keterbatasan kemampuan penulis. Oleh karena itu, penulis mengharapkan kritik dan saran yang sifatnya membangun dari pembaca untuk penyempurnaan isi Tugas Akhir ini.

Medan, 12 Maret 2025

Penulis



HARYA SYAFNELDI

NPM : 188160105

DAFTAR ISI

HALAMAN PENGESAHAN	i
HALAMAN PERNYATAAN	iii
HALAMAN PERNYATAAN PUBLIKASI	iii
RIWAYAT HIDUP	iv
ABSTRAK	v
<i>ABSTRACT</i>	vi
KATA PENGANTAR	vii
DAFTAR ISI	viii
DAFTAR GAMBAR	x
DAFTAR TABEL	xii
1. BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Batasan Masalah	4
1.3 Tujuan Penelitian	4
2. BAB II LANDASAN TEORI	6
2.1 Principal Component Analysis (PCA)	6
2.2 Tanaman Jamur	7
2.3 Data Mining	7
2.4 Klasifikasi	14
2.5 K - Nearest Neighbor	14
2.6 Confusion Matrix	16
2.7 Flowchart	19
3. BAB III METODOLOGI PENELITIAN	21
3.1 Proses Penelitian	21
3.2 Metode Penelitian	22
3.3 Data Yang Digunakan	22
3.4 Analisis Kinerja Metode KNN dan PCA+KNN	26
4. BAB IV HASIL DAN PEMBAHASAN	27
4.1 Preprocessing Data	27
4.2 Analisis Principal Component Analysis (PCA)	30
4.3 Confusion Matrix untuk Uji Akurasi	38
4.4 K-Nearest Neighbor (KNN)	39
4.5 <i>K-Nearest Neighbor</i> (KNN) dengan PCA	42
4.6 Pembahasan	45

5. BAB V KESIMPULAN.....	47
5.1 Kesimpulan.....	47
5.2 Saran	47
6. DAFTAR PUSTAKA.....	48



DAFTAR GAMBAR

	Halaman
Gambar 3.1 Proses Peningkatan Akurasi KNN menggunakan PCA	21
Gambar 3.2 Dataset Jamur.....	23
Gambar 3.3 Flowchart Sistem yang berjalan.....	25
Gambar 3.4 Proses Penyelesaian Masalah.....	26
Gambar 4.1 Hasil Matriks Kovarians.....	32
Gambar 4.2 Output.....	33
Gambar 4.3 Visualisasi Data Jamur Setelah PCA	37
Gambar 4.4 Inisialisasi Kode KNN	39
Gambar 4.5 Confusion Matrix KNN.....	40
Gambar 4.6 Confusion Matrix KNN dengan PCA	42
Gambar 4.7 Perbandingan Akurasi KNN dan KNN dengan PCA	43



DAFTAR TABEL

	Halaman
Tabel 2.1 Simbol Use Case Diagram	17
Tabel 2.2 Simbol Activity Diagram	18
Tabel 2.3 Simbol-simbol FlowChart.....	19
Tabel 3.1 Jamur tiram dataset	24
Tabel 4.1 Tipe Dataset.....	27
Tabel 4.2 Hasil Normalisasi.....	29
Tabel 4.3 Hasil Dekomposisi.....	34
Tabel 4.4 Confusion Matrix.....	38
Tabel 4.5 Hasil Prediksi KNN	39
Tabel 4.6 Hasil Prediksi Model.....	44



BAB I

PENDAHULUAN

1.1 Latar Belakang

Pengenalan dan klasifikasi spesies jamur merupakan salah satu tantangan dalam bidang biologi dan ekologi. Jamur memiliki peran penting dalam ekosistem, baik sebagai pengurai maupun sebagai sumber makanan bagi berbagai organisme. Namun, klasifikasi jamur seringkali sulit dilakukan karena keragaman bentuk, warna, dan tekstur yang sangat bervariasi. Untuk mengatasi masalah ini, teknologi data mining dan machine learning telah digunakan untuk mengembangkan model klasifikasi yang dapat membantu identifikasi jenis-jenis jamur secara lebih akurat.

Salah satu algoritma yang umum digunakan dalam klasifikasi adalah K-Nearest Neighbors (KNN). KNN merupakan algoritma sederhana yang efektif dalam mengklasifikasikan data berdasarkan kedekatan jarak antara data yang diuji dan data latih. Meskipun KNN mudah diimplementasikan dan memberikan hasil yang cukup baik, akurasi sering kali terpengaruh oleh dimensi data yang tinggi. Dalam konteks klasifikasi jamur, data yang dihasilkan dari pengukuran morfologis dan kimiawi jamur dapat memiliki ratusan fitur. Hal ini menyebabkan masalah "*curse of dimensionality*," yang dapat mengakibatkan penurunan akurasi klasifikasi.

Untuk meningkatkan akurasi KNN, salah satu metode yang dapat diterapkan adalah Principal Component Analysis (PCA). PCA adalah teknik reduksi dimensi yang berfungsi untuk mengurangi jumlah fitur dalam dataset dengan tetap mempertahankan informasi yang paling penting. Dengan

menggunakan PCA, fitur yang kurang relevan dapat dihilangkan, sehingga data menjadi lebih ringkas dan lebih mudah untuk dianalisis. Melalui proses ini, PCA dapat membantu meningkatkan akurasi model KNN dengan mengurangi noise dan mengoptimalkan ruang pencarian untuk klasifikasi.

Penelitian ini bertujuan untuk menginvestigasi pengaruh penerapan PCA terhadap akurasi KNN dalam klasifikasi data jenis jamur. Dengan melakukan eksperimen menggunakan dataset jamur yang sudah tersedia, penelitian ini diharapkan dapat memberikan wawasan baru tentang cara efektif untuk meningkatkan akurasi model klasifikasi jamur, serta berkontribusi pada pengembangan teknik pengenalan spesies jamur yang lebih baik.

Klasifikasi memang memiliki peran penting dalam machine learning dan data mining. Ketika kita memiliki data yang belum memiliki label kelas, kita dapat menggunakan algoritma klasifikasi untuk memprediksi kelas dari objek-objek tersebut. Klasifikasi adalah pola yang dilakukan secara terarah. (Tuntun, 2022). Algoritma klasifikasi, seperti *Naive Bayes*, *Decision Trees*, atau *Support Vector Machines*, mempelajari pola dari data yang telah diberi label dan kemudian dapat memprediksi kelas dari data yang belum diberi label. Ini memungkinkan untuk mengambil keputusan berdasarkan informasi yang ada.

Dalam dunia data mining, klasifikasi merupakan salah satu tugas utama yang bertujuan untuk memprediksi label kategori dari objek berdasarkan fitur-fiturnya. Klasifikasi membantu dalam membedakan antara objek yang satu dengan yang lainnya, yang sangat berguna dalam berbagai aplikasi seperti pengenalan pola, pengambilan keputusan, dan lain-lain. Salah satu metode klasifikasi yang paling dasar namun kinerjanya efektif adalah K-Nearest Neighbor

(KNN). Metode ini melakukan klasifikasi terhadap objek dengan mempertimbangkan 'tetangga' terdekatnya dalam ruang fitur. KNN adalah metode yang intuitif dan mudah diimplementasikan, di manaklasifikasi dilakukan dengan mengukur jarak antara titik data baru dengan titik data yang telah diklasifikasikan sebelumnya.

Studi yang dilakukan oleh Dhany (2021) mengevaluasi dataset Penyakit Jantung dengan menerapkan metode *K-Nearest Neighbor* pada *confusion matrix* untuk menguji akurasi. Hasilnya menunjukkan tingkat akurasi sebesar 81.31% *Classification Error* pada klasifikasi tertentu, menandakan bahwa dataset ini dapat dijadikan referensi untuk penelitian mendatang dalam upaya meningkatkan akurasi prediksi penyakit jantung.

Dalam penelitian Dhany (2022), identifikasi jenis data jamur dilakukan dengan menerapkan metode *K-Nearest Neighbor*. Klasifikasi pada data Jamur Tiram menggunakan nilai ($k = 5$) menghasilkan skor akurasi sebesar 71.42%. Penelitian ini juga menemukan bahwa jumlah data latih dan data uji yang digunakan berpengaruh terhadap skor akurasi yang diperoleh.

Kelemahan pada metode klasifikasi tersebut dapat mempengaruhi *performance* dalam mengelompokkan data. *Performance* dapat diartikan sebagai tingkat pencapaian hasil. Oleh karena itu berdasarkan latar belakang yang telah dibahas sebelumnya, maka penelitian ini dilakukan untuk meningkatkan *performance* metode *K-NN* melalui pembobotan terhadap model jarak (*similarity*) menggunakan *Principal Component Anaysis* (PCA), diharapkan hal ini mampu mengatasi kelemahan pada *K-NN* dan meningkatkan performa dari *KNN* dalam mengklasifikasi terhadap data yang digunakan.

1.2 Rumusan Masalah

Salah satu kekurangan dari metode klasifikasi K-NN terletak pada akurasi yang belum maksimal. Hal ini disebabkan oleh adanya sejumlah atribut yang kurang relevan, yang berdampak pada hasil klasifikasi. Oleh karena itu, diperlukan suatu pendekatan untuk menyaring atau mengurangi atribut-atribut yang tidak relevan guna meningkatkan efektivitas perhitungan jarak dan pada akhirnya dapat memperbaiki akurasi klasifikasi menggunakan metode K-NN.

1.3 Batasan Masalah

Penelitian ini memiliki beberapa ruang lingkup pembahasan, di antaranya:

1. Fokus pada penerapan metode seleksi fitur menggunakan Principal Component Analysis (PCA) dalam upaya meningkatkan akurasi klasifikasi dengan algoritma K-Nearest Neighbor (K-NN). Metode klasifikasi yang digunakan dalam penelitian ini adalah K-NN (*K-Nearest Neighbour*)
2. Analisis kinerja dari metode klasifikasi K-NN (*K-Nearest Neighbour*) berdasarkan pengukuran akurasi menggunakan *Confusion Matrix*.
3. Penelitian ini menggunakan 159 dataset jamur.

1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah untuk mereduksi dan menyaring atribut-atribut yang kurang relevan tanpa menghilangkan makna serta informasi penting dari data asli, dengan memanfaatkan metode Principal Component Analysis (PCA) guna meningkatkan akurasi dalam proses klasifikasi menggunakan algoritma K-Nearest Neighbor (K-NN).

1.4. Manfaat Penelitian

Manfaat yang diharapkan dari penelitian ini antara lain sebagai berikut:

1. Algoritma K-Nearest Neighbor (K-NN) memiliki karakteristik yang sederhana dan mudah diaplikasikan, sehingga cocok digunakan oleh pemula dalam bidang data science dan machine learning.
2. Penelitian ini bertujuan untuk meningkatkan tingkat akurasi klasifikasi pada algoritma K-Nearest Neighbor (K-NN) melalui penerapan metode Principal Component Analysis (PCA), serta melakukan evaluasi kinerja dengan membandingkannya terhadap penerapan K-NN tanpa reduksi fitur, berdasarkan hasil akurasi yang diperoleh.
3. Memberikan wawasan yang lebih mendalam mengenai efektivitas metode seleksi fitur menggunakan Principal Component Analysis (PCA) dalam meningkatkan performa klasifikasi K-Nearest Neighbor (K-NN).

BAB II LANDASAN TEORI

2.1 Principal Component Analysis (PCA)

Principal Component Analysis (PCA) merupakan metode reduksi dimensi secara linear yang sering digunakan dalam eksplorasi data, visualisasi, serta tahap pra-pemrosesan. Teknik ini mentransformasikan data ke dalam sistem koordinat baru, di mana arah komponen utama yang merepresentasikan variasi terbesar dalam data dapat diidentifikasi dengan lebih mudah.

Komponen utama dari kumpulan titik dalam ruang koordinat nyata adalah serangkaian vektor satuan, di mana vektor ke- (n) adalah arah dari garis yang paling cocok dengan data sambil tetap ortogonal terhadap vektor sebelumnya. Garis yang paling cocok didefinisikan sebagai garis yang meminimalkan jarak tegak lurus rata-rata kuadrat dari titik-titik ke garis. Arah ini (yaitu, komponen utama) merupakan basis ortonormal di mana dimensi individu yang berbeda dari data adalah tidak berkorelasi secara linear.

Banyak studi menggunakan dua komponen utama pertama untuk memplot data dalam dua dimensi dan untuk mengidentifikasi secara visual kluster titik data yang saling berhubungan erat. PCA memiliki aplikasi di banyak bidang seperti genetika populasi, studi mikrobiom, dan ilmu atmosfer.

Secara umum, PCA digunakan ketika banyak variabel sangat berkorelasi satu sama lain dan diinginkan untuk mengurangi jumlahnya menjadi satu set yang independen. Komponen utama pertama dapat didefinisikan sebagai arah yang memaksimalkan varians dari data yang diproyeksikan. Komponen utama ke- (n) dapat diambil sebagai arah ortogonal terhadap komponen utama pertama yang

memaksimalkan varians dari data yang diproyeksikan. Untuk kedua tujuan tersebut, dapat ditunjukkan bahwa komponen utama adalah vektor *eigen* dari matriks kovariansi data. Dengan demikian, komponen utama sering dihitung dengan dekomposisi *eigen* dari matriks kovariansi data atau dekomposisi nilai tunggal dari matriks data.

2.2 Tanaman Jamur

Jamur merupakan organisme eukariotik heterotrof yang memerlukan senyawa organik sebagai sumber nutrisinya, baik dalam bentuk uniseluler maupun multiseluler (Hyde, 2024). Organisme ini biasanya hidup sebagai saprofit atau parasit untuk memenuhi kebutuhan makanannya. Dalam kehidupan sehari-hari, jamur dikenal dengan berbagai nama seperti kapang, khamir, ragi, dan cendawan. Organisme ini memiliki spora berding yang mengandung kitin, tidak mengandung plastid, tidak bersifat fagotrof, serta tidak mampu melakukan fotosintesis. Umumnya jamur memiliki struktur hifa berding, dengan satu atau banyak inti (mononukleat atau multinukleat), dan memperoleh nutrisi melalui proses absorpsi. Karakteristik jamur membedakannya dari organisme lain dalam hal struktur tubuh, cara makan, reproduksi, dan pertumbuhan.

Dari segi ukuran, jamur dibedakan menjadi dua kategori: mikroskopis dan makroskopis. Jamur makroskopis, seperti jamur merang, jamur kuping, dan jamur tempe, dapat dilihat langsung tanpa alat bantu. Sementara itu, jamur mikroskopis hanya dapat diamati melalui mikroskop. Bentuk tubuh jamur sangat bervariasi, mulai dari oval, bulat, berbentuk payung, mangkuk, hingga menyerupai kancing dengan bercak di bagian kepala.

Berdasarkan struktur dan cara reproduksinya, jamur dikelompokkan menjadi lima divisi (Bahar et al., 2022):

1. Chytridiomycota

Merupakan jamur tingkat rendah yang dianggap sebagai bentuk transisi antara protista dan fungi. Umumnya hidup di tanah dan lingkungan perairan, dengan sekitar 1000 spesies yang telah diidentifikasi.

2. Zygomycota

Termasuk jamur tingkat rendah yang berkembang biak secara aseksual melalui pembentukan zigospora. Sekitar 1000 spesies dikenal dan sering ditemukan pada makanan yang telah basi.

3. Ascomycota

Berkembang biak secara seksual dengan menghasilkan askospora dan memiliki hifa yang bersepta. Terdapat sekitar 65.000 spesies yang hidup di lingkungan darat, air tawar, dan laut. Genus penting dalam kelompok ini meliputi *Aspergillus*, *Penicillium*, dan *Saccharomyces*.

4. Basidiomycota

Memiliki struktur tubuh buah lengkap seperti tudung, bilah (lamela), batang, cincin, dan cawan. Terdiri dari benang-benang panjang yang disebut hifa, yang menyebar membentuk miselium guna menyerap nutrisi. Sekitar 30.000 spesies diketahui, termasuk *Volvariella volvacea* (jamur merang) yang dapat dikonsumsi.

5. Deuteromycota

Merupakan kelompok jamur yang tidak melalui fase reproduksi seksual, sehingga disebut juga fungi imperfecti. Meskipun demikian, mereka berkembang biak

secara aseksual dan memiliki hifa bersekat. Contoh spesies dalam divisi ini adalah *Epidermophyton floccosum*, penyebab kutu air.

2.2.1. Karakteristik Jamur Beracun

Selain memiliki banyak manfaat, jamur juga dapat membawa dampak negatif. Jamur telah digunakan dalam berbagai industri, seperti pembuatan roti, keju, bir, serta produksi antibiotik. Namun, beberapa jenis jamur dapat membahayakan kesehatan manusia, dengan ciri-ciri sebagai berikut:

1. Memiliki struktur unik seperti cincin atau cawan pada batang, serta bintik-bintik pada bagian bawah tudung.
2. Mempunyai warna mencolok seperti merah, hijau, biru, atau hitam, meskipun ada juga jamur berwarna gelap yang aman dikonsumsi.

Dampak Negatif Mengonsumsi Jamur Beracun:

1. Halusinasi

Beberapa jenis jamur mengandung zat toksik yang dapat menimbulkan efek halusinogen, sakit kepala, gangguan otot, serta peningkatan detak jantung, seperti jamur dari genus *Galerina*.

2. Gangguan Pencernaan

Menyebabkan mual, muntah, kram perut, dan diare. Salah satu contohnya adalah *Amanita phalloides*.

3. Kerusakan Hati

Zat beracun dalam *Amanita phalloides* atau *Amanita muscaria* dapat menyebabkan gagal hati dalam 6–24 jam setelah dikonsumsi, disertai gejala awal berupa muntah dan nyeri perut, yang berlanjut pada gangguan ginjal

4. Kematian

Konsumsi jamur beracun dalam waktu kurang dari dua hari dapat menyebabkan kematian.

2.2.2. *Potensi Ekonomi Jamur*

Jamur merupakan bahan pangan bergizi tinggi dan bernilai ekonomi. Indonesia sebagai negara agraris memiliki potensi besar dalam pengembangan berbagai jenis tanaman, termasuk jamur, sebagai peluang usaha di sektor agribisnis. Hal ini membuka jalan bagi masyarakat untuk menciptakan usaha mandiri dan inovatif.

Menurut Jaelani (2008), jamur tiram mengandung berbagai mineral penting seperti zat besi (Fe), fosfor (P), natrium (Na), kalium (K), dan kalsium (Ca). Selain itu, jamur tiram juga mengandung polisakarida yang berperan dalam meningkatkan sistem imun tubuh.

Sebagai bagian dari kingdom Fungi, jamur tidak memiliki struktur daun atau akar sejati dan tidak mengandung klorofil, sehingga tidak bisa berfotosintesis seperti tumbuhan pada umumnya. Karena perbedaan ini, jamur diklasifikasikan tersendiri, terpisah dari kelompok tumbuhan dan hewan. Jamur ada yang makroskopis dan ada pula yang mikroskopis. Kebanyakan jamur bersifat multiseluler seperti jamur tempe dan merang, namun ada juga yang uniseluler seperti ragi (*Saccharomyces*). Struktur tubuh jamur multiseluler tersusun dari benang halus bernama hifa. Di bawah mikroskop, hifa bisa tampak bersekat (berseptata) atau tidak bersekat (Gudang Biologi, 2015).

2.3. **Data Mining**

Data mining merupakan suatu teknologi yang mengintegrasikan metode analisis tradisional dengan algoritma canggih untuk mempercepat pemrosesan

data dalam skala besar. Istilah ini sering digunakan untuk merujuk pada proses pencarian pengetahuan tersembunyi dalam basis data. Proses ini memanfaatkan pendekatan dari bidang statistik, matematika, kecerdasan buatan, dan pembelajaran mesin guna mengeksplorasi dan menganalisis data dalam jumlah besar (Kumar et al., 2020).

Kemampuan analitis data mining melampaui sistem pendukung keputusan konvensional, yang cenderung memerlukan lebih banyak waktu dan biaya. Dengan menelusuri basis data secara mendalam, data mining dapat menemukan pola tersembunyi dan informasi yang berguna—yang mungkin tidak disadari oleh pelaku bisnis.

Seiring kemajuan teknologi, pengumpulan dan penyimpanan data dalam jumlah besar menjadi semakin mudah. Dalam konteks ini, data mining hadir sebagai alat yang dapat secara otomatis mengekstrak informasi berharga dari data yang tersimpan. Istilah lain yang sering digunakan secara bergantian adalah Knowledge Discovery in Databases (KDD), yang merujuk pada proses penemuan pengetahuan tersembunyi dalam database besar, meskipun kedua istilah ini memiliki makna yang sedikit berbeda namun saling berkaitan.

1. Teknik-Teknik Data Mining

Data mining bertujuan untuk menemukan nilai tambah dari data besar yang tidak dapat dianalisis secara manual. Proses ini memungkinkan analisis otomatis terhadap data berskala besar guna mengidentifikasi pola penting yang

tidak tampak jelas, serta mengekstrak informasi tersembunyi. Teknik-teknik utama dalam data mining antara lain:

a. Klasifikasi (*Classification*)

Teknik ini digunakan untuk mengkategorikan data berdasarkan perilaku atau atribut yang telah diketahui sebelumnya. Pendekatan ini bersifat supervised learning, di mana data dilatih menggunakan sampel data yang sudah diklasifikasikan untuk membuat model prediksi. Salah satu metode populer adalah *decision tree*, yang merepresentasikan proses klasifikasi melalui struktur pohon keputusan.

b. Asosiasi (*Association*)

Teknik ini bertujuan menemukan hubungan atau keterkaitan antara item dalam suatu dataset. Misalnya, analisis pasar swalayan dapat menunjukkan bahwa pelanggan yang membeli roti juga cenderung membeli susu. Tingkat relevansi aturan asosiasi ini diukur menggunakan parameter *support* dan *confidence*.

c. Pengelompokan (*Clustering*)

Berbeda dengan klasifikasi, clustering tidak memerlukan label data sebelumnya. Teknik ini mengelompokkan data berdasarkan kesamaan atribut dan biasanya menggunakan metode statistik atau neural network. Tujuannya adalah memaksimalkan kesamaan dalam kelompok dan meminimalkan kesamaan antar kelompok.

2. Tahapan Proses Data Mining (KDD)

Dalam praktiknya, penerapan data mining pada dataset skala besar memerlukan pendekatan yang sistematis. Proses ini tidak hanya mencakup analisis, tetapi juga mencakup persiapan data dan interpretasi hasil. Berikut ini adalah lima tahapan utama dalam proses Knowledge Discovery in Databases (KDD) menurut Han (2019):

1. Seleksi Data (Data Selection)

Pemilihan data yang relevan dari sistem operasional untuk diproses dalam tahapan selanjutnya. Data ini disimpan secara terpisah dari database utama.

2. Pembersihan Data (Pre-processing/Cleaning)

Tahap ini mencakup penghapusan data duplikat, perbaikan inkonsistensi, dan koreksi kesalahan data. Proses *enrichment* juga dapat dilakukan untuk melengkapi data dengan informasi tambahan.

3. Transformasi Data (Transformation/Coding)

Pada tahap ini, data diubah ke dalam format yang sesuai agar dapat dianalisis secara efektif. Proses ini tergantung pada pola yang ingin ditemukan.

4. Penerapan Data Mining

Penggunaan algoritma tertentu untuk menemukan pola atau informasi menarik dalam dataset. Pemilihan metode tergantung pada tujuan analisis.

5. Interpretasi dan Evaluasi

Pola yang dihasilkan dari proses data mining disajikan dalam bentuk yang mudah dipahami. Tahap ini juga mencakup evaluasi terhadap kesesuaian pola dengan pengetahuan atau hipotesis yang telah ada. Proses KDD bersifat iteratif, memungkinkan analisis untuk kembali ke tahap sebelumnya jika hasil yang diperoleh belum memuaskan.

2.3 Klasifikasi

Klasifikasi merupakan proses yang umumnya digunakan untuk memprediksi kategori atau kelas dari suatu data, serta menyusun data ke dalam kelompok tertentu berdasarkan model yang telah dibangun melalui proses pelatihan data sebelumnya. Model ini digunakan untuk menetapkan nilai-nilai kelas berdasarkan atribut yang dimiliki, dan kemudian diaplikasikan pada data baru yang belum diketahui kelasnya. Teknik klasifikasi banyak dimanfaatkan dalam berbagai bidang seperti evaluasi kelayakan kredit, pemasaran yang ditargetkan, diagnosa medis, serta penilaian efektivitas suatu kebijakan. Proses klasifikasi melibatkan identifikasi sekumpulan kelas yang telah ditentukan sebelumnya dan membangun model yang mampu mengklasifikasikan data (*tuple*) yang belum memiliki label kelas. Model tersebut dapat direpresentasikan dalam bentuk aturan klasifikasi, struktur pohon keputusan, atau rumus matematika. Beberapa metode klasifikasi yang umum digunakan antara lain adalah *Decision Tree*, *Bayesian Network*, *Adaptive Bayesian Network*, *Naïve Bayes*, dan lainnya.

2.4 K - Nearest Neighbor

Salah satu metode yang termasuk dalam kelompok pembelajaran berbasis instance adalah algoritma *K-Nearest Neighbor* (KNN). Teknik ini bekerja dengan

mencari sejumlah k objek dari data pelatihan yang memiliki jarak paling dekat dengan data uji atau data baru yang ingin diklasifikasikan. KNN digunakan untuk melakukan klasifikasi terhadap objek dengan mengacu pada data pelatihan yang memiliki kedekatan jarak dengan data yang diuji. Pendekatan ini didasarkan pada perbandingan antara permasalahan baru dengan kasus-kasus sebelumnya, dengan menghitung bobot berdasarkan penjumlahan fitur yang dimiliki.

Dalam metode ini, selama proses pelatihan, data disimpan dalam bentuk *vector* fitur beserta label klasifikasinya. Ketika sebuah data baru akan diklasifikasikan, algoritma akan mengukur jarak antara vektor tersebut dengan vektor-vektor dalam data pelatihan dan memilih k vektor yang memiliki jarak terdekat.

Langkah-langkah dalam proses perhitungan KNN antara lain:

- a. Menentukan nilai k sebagai jumlah tetangga terdekat yang akan dipertimbangkan.
- b. Menghitung jarak Euclidean antara data uji dan setiap data dalam dataset pelatihan.
- c. Mengurutkan data berdasarkan jarak terkecil ke terbesar.
- d. Mengumpulkan label kategori (Y) dari data terdekat.
- e. Menentukan prediksi berdasarkan mayoritas label dari k tetangga terdekat tersebut.

2.4.1 Model Pengukuran Jarak

Metode ini digunakan untuk mengukur tingkat kemiripan antar data. Setiap model pengukuran jarak memiliki keunggulannya masing-masing. Misalnya, *Manhattan distance* dikenal efektif dalam mendeteksi outlier, sementara *Euclidean distance* menghitung jarak lurus antara dua titik, yang cocok digunakan untuk data yang terdistribusi secara merata. Namun, Euclidean terkadang kurang akurat dalam merepresentasikan kecocokan antar data (Abdalla, 2022).

Dalam KNN, pengukuran kesamaan antar data dilakukan dengan menggunakan model jarak. Akan tetapi, penggunaan model jarak konvensional memiliki kelemahan, terutama ketika terdapat atribut yang nilainya sangat besar sehingga dapat mendominasi atribut lain dan menurunkan akurasi prediksi. Masalah ini menjadi lebih signifikan ketika menangani data dengan jumlah sampel yang sangat besar (Dhany, 2021).

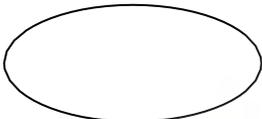
2.5 Confusion Matrix

Pengujian klasifikasi menggunakan metode KNN terhadap data yang diuji menghasilkan proses perhitungan, yang kemudian dituangkan ke dalam tabel yang dikenal sebagai confusion matrix (Dhany, 2021). Confusion matrix menggambarkan parameter keberhasilan dan kesalahan dalam proses klasifikasi data uji ke dalam dua kelas berbeda, yaitu kelas positif dan kelas negatif.

2.7.1. Use Case Diagram

Use case diagram adalah bentuk pemodelan yang menggambarkan perilaku dari sistem informasi yang akan dibangun. Diagram ini digunakan untuk mengidentifikasi seluruh fungsi dalam sistem, serta menentukan siapa saja pengguna yang berhak mengakses fungsi-fungsi tersebut.

Tabel 2.1 Simbol Use Case Diagram

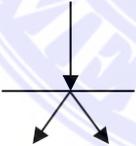
Simbol	Keterangan
 User Case	Menggambarkan fitur atau layanan yang disediakan sistem dalam bentuk unit yang berinteraksi dengan aktor.
 Actor	Representasi dari pengguna atau sistem lain yang memanfaatkan fitur sistem.
 Asosiasi	Garis tanpa panah yang menunjukkan hubungan langsung antara aktor dan use case.
 Generalization	Panah terbuka yang menunjukkan aktor berperan secara pasif.
 include	Use case bagian wajib dari use case lain.
 extends	Perluasan use case jika kondisi tertentu terpenuhi.

(Sumber: Ansori, 2020)

2.7.2. Activity Diagram

Activity diagram menampilkan alur kerja atau aktivitas dari suatu sistem maupun proses bisnis. Simbol-simbol yang digunakan untuk menggambarannya dijelaskan dalam Tabel

Tabel 2.2 Simbol Activity Diagram

Simbol	Keterangan
 <i>Start Point</i>	Titik awal aktivitas.
 <i>End Point</i>	Akhir dari aktivitas.
 <i>Activities</i>	Mewakili proses dalam sistem.
 Fork/percabangan	Menunjukkan aktivitas paralel atau penggabungan.
 Join (penggabungan)	Menunjukkan proses dekomposisi atau penggabungan alur.

 <i>Decision point</i>	Simbol pengambilan keputusan (true/false).
 <i>Swimlane</i>	Pembagian aktivitas berdasarkan pelaku.

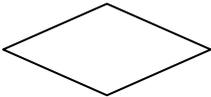
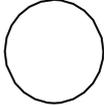
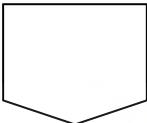
(Sumber: Ansori, 2020)

2.6 Flowchart

Flowchart menggambarkan urutan langkah-langkah dalam suatu proses menggunakan simbol-simbol standar yang tersusun secara sistematis (Manurung, 2020). Menurut Manurung (2020), flowchart adalah representasi grafis dari prosedur atau langkah dalam program. Flowchart membantu analis dan programmer memecah masalah menjadi bagian-bagian kecil serta mengevaluasi alternatif penyelesaian. Flowchart memudahkan pemahaman masalah yang kompleks dan digunakan untuk merancang maupun mendesain program. Oleh karena itu, diagram ini harus mampu mewakili berbagai elemen dalam bahasa pemrograman.

Tabel 2.3 Simbol-simbol FlowChart

No.	Simbol	Keterangan
1.	 <i>Predefined Process</i>	Permulaan suatu subprogram.

2.	 Keputusan	Menyatakan perbandingan atau seleksi data.
3.	 Penghubung	Menyambungkan bagian-bagian dalam satu halaman.
4.	 <i>Offline Connector</i>	Menghubungkan bagian di halaman berbeda.
5.	 Terminator	Menunjukkan awal atau akhir program.
6.	 Anak Panah	Menyatakan arah alur program.
7.	 Preparation	Proses inisialisasi atau pemberian nilai awal.
8.	 Process	Menggambarkan proses pengolahan data.
9.	 <i>Input/Output</i>	Proses input atau output data.

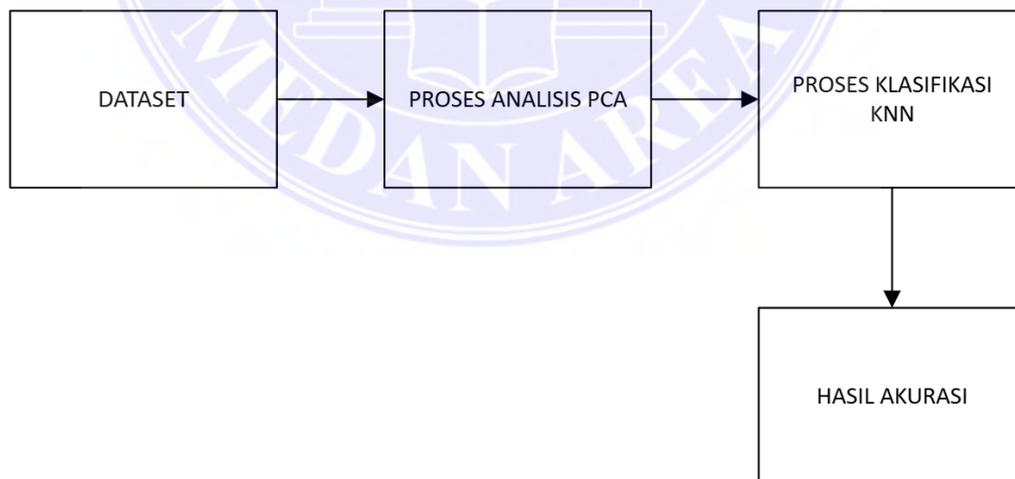
(Sumber: Ansori, 2020)

BAB III METODOLOGI PENELITIAN

3.1 Proses Penelitian

Penelitian ini berusaha meningkatkan akurasi pengklasifikasian *K-Nearest Network* (KNN) menggunakan metode *Principal Component Analysis* (PCA) dengan cara menyederhanakan data tanpa mengurangi maksud dan tujuan dari data tersebut sehingga diharapkan atribut-atribut yang kurang relevan dapat dikurangi pengaruhnya terhadap proses pengklasifikasian menggunakan *K-Nearest Network* (KNN).

Untuk memberikan gambaran yang lebih jelas mengenai proses dalam penelitian ini, subbab berikut akan menguraikan setiap tahapannya secara bertahap. Secara umum, alur dari tahapan-tahapan tersebut ditampilkan pada Gambar 3.1.



Gambar 3.1 Proses Peningkatan Akurasi KNN menggunakan PCA

Berdasarkan gambar 3.1 terlihat bahwa untuk meningkatkan akurasi

pengklasifikasian *K-Nearest Network* menggunakan metode *Principal Component Analysis* (PCA), adapun tahapan-tahapannya adalah :

- 1 *Data Preprocessing*
- 2 Lakukan proses analisis PCA
- 3 Lakukan proses klasifikasi *K-Nearest Network*
- 4 Analisis pengujian model klasifikasi menggunakan tabulasi *Confusion Matrix* (*two-class prediction*)

3.2 Metode Penelitian

Langkah-langkah yang akan ditempuh dalam pelaksanaan penelitian ini meliputi:

1. Kajian Literatur

Tahap awal berupa kajian literatur dilakukan dengan mengumpulkan berbagai referensi yang berkaitan dengan algoritma K-NN yang diintegrasikan dengan *Principal Component Analysis* (PCA). Sumber-sumber yang digunakan berasal dari buku, jurnal ilmiah, artikel, serta referensi relevan lainnya.

2. Perancangan Sistem

Pada tahap ini, dilakukan perancangan arsitektur sistem, pengumpulan data pelatihan, serta perancangan antarmuka pengguna. Seluruh proses perancangan didasarkan pada hasil kajian literatur yang telah dianalisis sebelumnya.

3. Tahap Pengujian

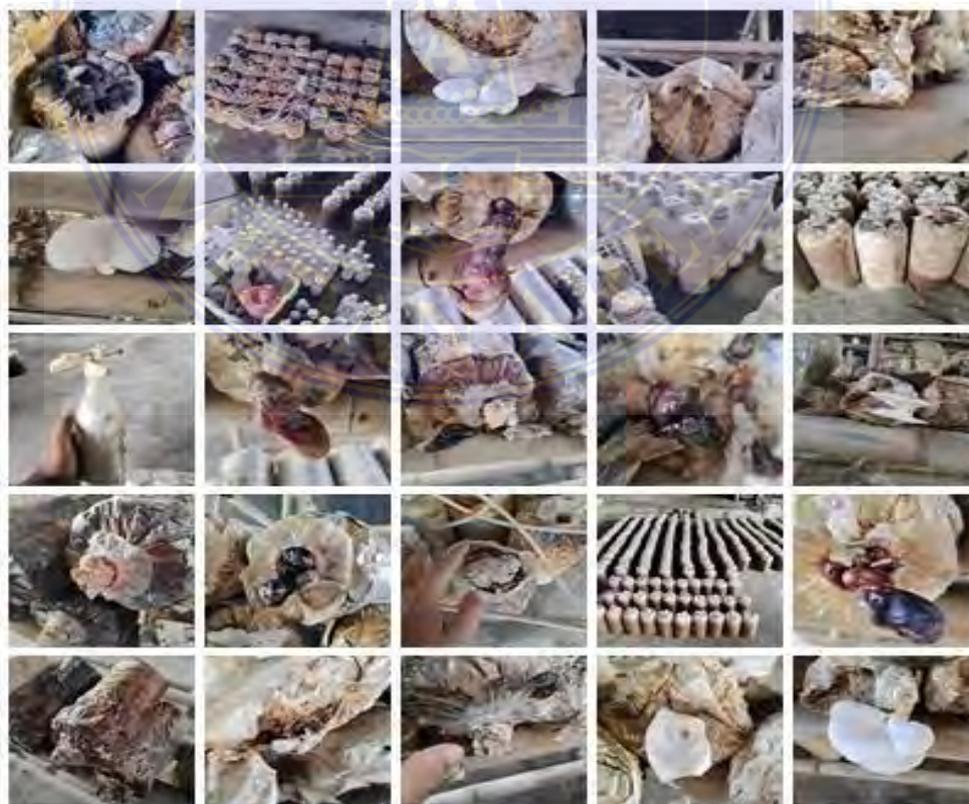
Proses pengujian dilakukan untuk memastikan perhitungan jarak dan pemilihan atribut dalam sistem telah berjalan sebagaimana yang direncanakan.

4. Penyusunan Laporan

Tahap akhir berupa dokumentasi seluruh hasil analisis terhadap tingkat akurasi data menggunakan algoritma K-NN yang dikombinasikan dengan metode Principal Component Analysis (PCA).

3.3 Data Yang Digunakan

Penelitian ini bertujuan untuk menganalisa kinerja dari metode K-NN sehingga dibutuhkan data observasi kemudian penulis memilih untuk menganalisa menggunakan dataset data observasi Jamur tiram. Pada penelitian ini dibutuhkan perbandingan terhadap data observasi agar dapat menganalisa hasil pengujian. Berikut adalah tabel jamur tiram dataset:



Gambar 3.2 Dataset Jamur

Tabel 3.1 Jamur tiram dataset

No	Kelas	Bentuk Tudung Lonceng	Bentuk Tudung Datar	Permukaan Tudung Beralur	Permukaan Tudung Halus	Warna Tudung Coklat	Warna Tudung Putih	Warna Tudung Kuning	Warna Batang Tidak Menghitam	Warna Batang Menghitam	Populasi Berkelompok	Populasi Banyak	Populasi Menyebar	Populasi Sedikit
1	p	0	0	0	1	1	0	0	0	1	0	0	1	0
2	e	0	0	0	1	0	0	1	0	1	0	1	0	0
3	e	1	0	0	1	0	1	0	0	1	0	1	0	0
4	p	0	0	0	0	0	1	0	0	1	0	0	1	0
5	e	0	0	0	1	0	0	0	1	0	0	0	0	0
6	e	0	0	0	0	0	0	1	0	1	0	1	0	0
7	e	1	0	0	1	0	1	0	0	1	0	1	0	0
8	e	1	0	0	0	0	1	0	0	1	0	0	1	0
9	p	0	0	0	0	0	1	0	0	1	0	0	0	0

159	e	1	0	0	0	0	1	0	0	1	0	0	1	0

Keterangan data:

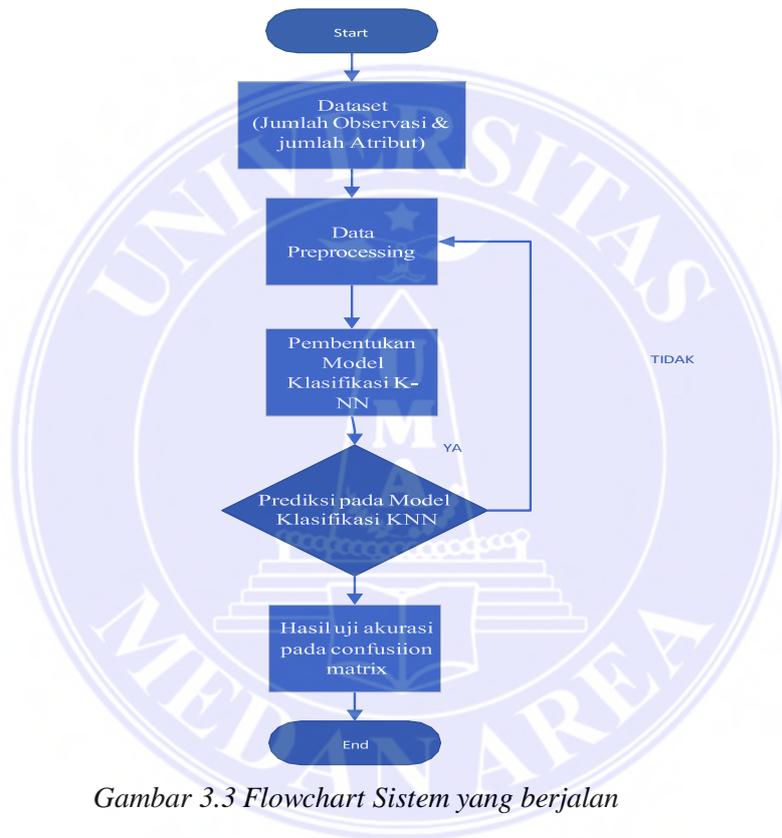
Kelas :

p = *Poisonous*

e = *Edible*

3.2. Analisis Sistem yang Sedang Berjalan

Analisis sistem berfungsi untuk memberikan gambaran mengenai sistem yang tengah digunakan saat ini. Diagram alir (flowchart) dari sistem berjalan merepresentasikan proses kerja dari sistem tersebut secara visual. Dalam penelitian ini, alur sistem yang sedang berlangsung disajikan melalui flowchart yang ditampilkan pada gambar di bawah ini.



Gambar 3.3 Flowchart Sistem yang berjalan

Pada Gambar 3.3, terlihat rancangan penelitian dibagi menjadi beberapa tahapan yakni data latih dan data uji. Setelah data di *preprocessing* langkah selanjutnya pada data latih dan uji adalah menghitung nilai proporsi varians kovariansnya, kemudian pada data uji data di klasifikasi menggunakan algoritma KNN. Selanjutnya adalah proses menghitung akurasi data menggunakan *Confusion Matrix*. Maka proses akhir adalah hasil akurasi dari algoritma KNN dan PCA.

3.4 Analisis Kinerja Metode KNN dan PCA+KNN

Untuk melihat apakah model penelitian yang diusulkan mampu meningkatkan performa pengklasifikasian KNN, maka akan dilakukan analisis kinerja PCA+KNN berdasarkan hasil tabulasi *Confusion Matrix (two-class prediction)* yang kemudian hasil tersebut akan dibandingkan dengan metode KNN konvensional.

Secara garis besar prosedur dari penelitian yang diusulkan dapat dilihat pada gambar 3.3



Gambar 3.4 Proses Penyelesaian Masalah

Berdasarkan Gambar 3.4, dapat dijelaskan alur proses penelitian adalah sebagai berikut :

1. Pemilihan Dataset, data yang digunakan terdiri dari satu set data Jamur
2. Dataset tersebut akan digunakan pada proses pengklasifikasian PCA+KNN dibandingkan dengan KNN konvensional
3. Selanjutnya dari setiap metode untuk setiap dataset akan dilakukan analisis pengujian model klasifikasi menggunakan tabulasi *Confusion Matrix (two-class prediction)*

BAB V KESIMPULAN

5.1 Kesimpulan

Penelitian ini berhasil menunjukkan bahwa integrasi *Principal Component Analysis* (PCA) dengan algoritma *K-Nearest Neighbors* (KNN) dapat meningkatkan performa klasifikasi pada dataset jamur. PCA berperan dalam mereduksi dimensi data dengan menghilangkan fitur-fitur yang kurang relevan. Dengan demikian, KNN dapat bekerja lebih efisien dengan fokus pada fitur-fitur paling informatif. Peningkatan akurasi yang signifikan dari 94.96% (KNN standar) menjadi 97.45% (KNN dengan PCA) menegaskan potensi PCA sebagai teknik pra-pemrosesan yang efektif untuk meningkatkan kinerja model klasifikasi.

5.2 Saran

1. Penelitian ini dapat diperluas dengan menerapkan PCA dan KNN pada dataset jamur yang lebih besar dan kompleks untuk menguji efektivitas kedua metode ini dalam skenario yang lebih menantang.
2. Selain KNN, PCA juga dapat dikombinasikan dengan algoritma klasifikasi lain. Hal ini akan memberikan wawasan tentang sejauh mana PCA dapat meningkatkan kinerja berbagai model klasifikasi.

DAFTAR PUSTAKA

- Abdalla, H. I., and Amer, A. A. (2022). "Towards highly-efficient k-nearest neighbor algorithm for big data classification," in 2022 5th International Conference on Networking, Information Systems and Security (New York City, NY: IEEE), 1–5.
- Aggarwal, C. C. (2021). *Data mining: The textbook*. Springer
- Ali, A., Hamraz, M., Kumam, P., Khan, D. M., Khalil, U., Sulaiman, M., et al. (2020). A k-nearest neighbours based ensemble via optimal model selection for regression. *IEEE Access* 8, 132095–132105. doi: 10.1109/ACCESS.2020.3010099
- Amira, K. (2020). Apa Itu Flowchart: Pengertian Menurut Ahli, Fungsi, dan Jenisnya. Diakses dari gamedia.com.
- Ansori, A. (2020, Maret 14). Pengertian Use Case Diagram: Tujuan, Fungsi, Simbol, dan Contohnya. Diakses dari ansoriweb.com.
- Bian, C., and Huang, G. Q. (2024). Air pollution concentration fuzzy evaluation based on evidence theory and the K-nearest neighbor algorithm. *Front. Environ. Sci.* 12:1243962. doi: 10.3389/fenvs.2024.1243962
- Chung, Y.-W., Khaki, B., Li, T., Chu, C., and Gadh, R. (2019). Ensemble machine learning-based algorithm for electric vehicle user behavior prediction. *Appl. Energy* 254, 113732. doi: 10.1016/j.apenergy.2019.113732
- Dhany, Hanna Willa. 2021. Performa Algoritma K-Nearest Neighbour dalam Memprediksi Penyakit Jantung. Prosiding SENATIKA 2021.
- El-Kenawy, E.-S. M., Mirjalili, S., Ghoneim, S. M., Eid, M. M., El-Said, M., Khan, Z. S., et al. (2021). Advanced ensemble model for solar radiation forecasting using sine cosine algorithm and Newton's laws. *IEEE Access* 9, 115750–115765. doi: 10.1109/ACCESS.2021.3106233

- Fowler, M. (2020). *UML Distilled: A Brief Guide to the Standard Object Modeling Language* (4th ed.). Addison-Wesley.
- Han, J., Pei, J., & Kamber, M. (2019). *Data mining: Concepts and techniques*. Elsevier
- Hyde, K.D., Baldrian, P., Chen, Y. *et al.* Current trends, limitations and future research in the fungi?. *Fungal Diversity* **125**, 1–71 (2024). <https://doi.org/10.1007/s13225-023-00532-5>
- Jafar, R., Awad, A., Hatem, I., Jafar, K., Awad, E., and Shahrour, I. (2023). Multiple linear regression and machine learning for predicting the drinking water quality index in Al-seine lake. *Smart Cities* 6, 2807–2827. doi: 10.3390/smartcities6050126
- Jiang, Z., Yang, S., Smith, P., and Pang, Q. (2023). Ensemble machine learning for modeling greenhouse gas emissions at different time scales from irrigated paddy fields. *Field Crops Res.* 292:108821. doi: 10.1016/j.fcr.2023.108821
- Kincai77. (2020, Oktober 24). Pengenalan UML: Sejarah dan Building Blocks (Hubungan, Diagram, Benda). Diakses dari blog.mycoding.id.
- Kumar, V., & Minz, S. (2020). Feature selection: A literature review. *Smart Computing Review*, 5(3), 211-229
- Kurita, T. (2020). *Principal Component Analysis (PCA)*. In: *Computer Vision*. Springer, Cham. https://doi.org/10.1007/978-3-030-03243-2_649-1
- Manurung, R. A. Y., & Manuputty, A. D. (2020). Perancangan Sistem Informasi Lembaga Kemahasiswaan Universitas Kristen Satya Wacana Salatiga. *Jurnal Sistem Informasi dan Teknologi*, Universitas Muria Kudus.
- Nugraha, Rezky & Hidayat, Eka & Shofa, Rahmi. (2023). Klasifikasi Jenis Buah Jambu Biji Menggunakan Algoritma Principal Component Analysis dan K-Nearest Neighbor. *Generation Journal*. 7. 1-7. 10.29407/gj.v7i1.17900.
- Ramesh Maheshwari, Nuclear behavior in fungal hyphae, *FEMS Microbiology Letters*, Volume 249, Issue 1, August 2005, Pages 7–14, <https://doi.org/10.1016/j.femsle.2005.06.031>

- Saadatfar, H., Khosravi, S., Joloudari, J. H., Mosavi, A., and Shamshirband, S. (2020). A new K-nearest neighbors classifier for big data based on efficient data pruning. *Mathematics* 8:286. doi: 10.3390/math8020286
- Sharma, S., and Lakshmi, L. R. (2023). "Improved k-NN regression model using random forests for air pollution prediction," in Presented at the International Conference on Smart Applications, Communications, and Networking (SmartNets) (Istanbul: SmartNets).
- Srisuradetchai P and Suksrikran K (2024) Random kernel k-nearest neighbors regression. *Front. Big Data* 7:1402384. doi: 10.3389/fdata.2024.1402384
- Srisuradetchai P and Suksrikran K (2024) Random kernel k-nearest neighbors regression. *Front. Big Data* 7:1402384. doi: 10.3389/fdata.2024.1402384
- Srisuradetchai, P. (2023). A novel interval forecast for k-nearest neighbor time series: a case study of durian export in Thailand. *IEEE Access*. 12, 2032–2044. doi: 10.1109/ACCESS.2023.3348078
- Syriopoulos, P.K., Kalampalikis, N.G., Kotsiantis, S.B. et al. kNN Classification: a review. *Ann Math Artif Intell* (2023). <https://doi.org/10.1007/s10472-023-09882-x>
- Tan, R., Ottewill, J. R., and Thornhill, N. F. (2020). Monitoring statistics and tuning of Kernel principal component analysis with radial basis function kernels. *IEEE Access* 8, 198328–198342. doi: 10.1109/ACCESS.2020.3034550
- Ukey, N., Yang, Z., Li, B., Zhang, G., and Hu, Y. Zhang, W. (2023). Survey on exact kNN queries over high-dimensional data space. *Sensors* 23:629. doi: 10.3390/s23020629
- Wong, T.-T., and Yeh, P.-Y. (2020). Reliable accuracy estimates from k-fold cross validation. *IEEE Trans. Knowl. Data Eng.* 32, 1586–1594. doi: 10.1109/TKDE.2019.2912815

Xiong, S. (2024). A unified framework of principal component analysis and factor analysis. Academy of Mathematics and Systems Science, Chinese Academy of Sciences Beijing, China.

Yao, Y., Li, Y., Jiang, B., and Chen, H. (2021). Multiple kernel k-means clustering by selecting representative kernels. *IEEE Trans. Neural Netw. Learn. Syst.* 32, 4983–4996. doi: 10.1109/TNNLS.2020.3026532

Yul Harry Bahar, Iwan Saskiawan, Gabriella Susilowati. Potensi Jamur Pangan sebagai Pangan Fungsional untuk Meningkatkan Daya Tahan Tubuh Manusia. *Jurnal Agroekoteknologi dan Agribisnis*, Vol 6 No 1 (Juni 2022): 45–58



PAPER NAME

HARYA SYAFNELDI.pdf

AUTHOR

HARYA SYAFNELDI

WORD COUNT

8280 Words

CHARACTER COUNT

49596 Characters

PAGE COUNT

52 Pages

FILE SIZE

1.2MB

SUBMISSION DATE

Jun 2, 2025 8:38 AM GMT+7

REPORT DATE

Jun 2, 2025 8:39 AM GMT+7

● **4% Overall Similarity**

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

- 3% Internet database
- 2% Publications database
- Crossref database
- Crossref Posted Content database

● **Excluded from Similarity Report**

- Submitted Works database
- Bibliographic material
- Cited material
- Abstract
- Small Matches (Less then 15 words)



UNIVERSITAS MEDAN AREA

FAKULTAS TEKNIK

Kampus I : Jalan Kolam Nomor 1 Medan Estate / Jalan Gedung PBSI, Medan 20223
Kampus II : Jalan Sei Serayu Nomor 70 A / Jalan Setia Budi Nomor 79 B, Medan 20112 Telepon : (061) 8225602, 8201994
Fax : (061) 8226331 HP : 0811 607 259 website: www.uma.ac.id Email : univ_medanarea@uma.ac.id

Nomor : 812/FT/01.10/III/2024
Lampiran : -
Hal : **Pembimbing Tugas Akhir**

19 Maret 2024

Yth. Pembimbing Tugas Akhir
Nurul Khairina S.Kom, M.Kom (Sebagai Pembimbing)
di Tempat

Dengan hormat, sehubungan telah dipenuhinya persyaratan untuk memperoleh Tugas Akhir dari mahasiswa atas :

Nama : HARYA SYAFNELDI
NIM : 188160105
Jurusan : TEKNIK INFORMATIKA

Maka dengan hormat kami mengharapkan kesediaan saudara :

Nurul Khairina S.Kom, M.Kom (Sebagai Pembimbing)

Adapun Tugas Akhir Skripsi berjudul :

"Peningkatan Akurasi KNN Dengan Menggunakan Metode Principal Component Analysis Pada Data Jenis Jamur."

SK Pembimbing ini berlaku selama enam bulan terhitung sejak SK ini diterbitkan. Jika proses pembimbing melebihi batas waktu yang telah ditetapkan, SK ini dapat ditinjau ulang.

Demikian kami sampaikan, atas kesediaan saudara diucapkan terima kasih.

Dekan,



Dr Eng. Supriatno.ST, MT.



UNIVERSITAS MEDAN AREA

FAKULTAS TEKNIK

Kampus I : Jalan Kolam Nomor 1 Medan Estate/Jalan PBSI Nomor 1 ☎ (061) 7366878, 7360168, 7364348, 7366781, Fax (061) 7366998 Medan 20223
Kampus II : Jalan Setiabudi Nomor 79 / Jalan Sei Serayu Nomor 70 A, ☎ (061) 8225902, Fax (061) 8225331 Medan 20122
 Website: www.teknik.uma.ac.id E-mail: uivv_medanarea@uma.ac.id

Nomor : 325 /FT.6/01.10/VII/2024

12 Juli 2024

Lamp : -

Hal :

Penelitian Dan Pengambilan Data Tugas Akhir

Yth. Pimpinan Jamur Tiram Binjai
 Jln. Wijaya Kesuma Gg. Nusa Indah No. 19
 Di
 Binjai

Dengan hormat,

Kami mohon kesediaan Bapak/Ibu berkenan untuk memberikan izin dan kesempatan kepada mahasiswa kami tersebut dibawah ini :

NO	NAMA	NPM	PRODI
1	Harya Syafneldi	188160105	Teknik Informatika

Untuk melaksanakan Penelitian dan Pengambilan Data Tugas Akhir pada perusahaan/Instansi yang Bapak/Ibu Pimpin.

Perlu kami jelaskan bahwa Pengambilan Data tersebut adalah semata-mata untuk tujuan ilmiah dan Skripsi yang merupakan salah satu syarat bagi mahasiswa tersebut untuk mengikuti ujian sarjana pada Fakultas Teknik Universitas Medan Area dan tidak untuk dipublikasikan, dengan judul penelitian :

Peningkatan Akurasi KNN dengan Menggunakan Metode *Principal Component Analysis* pada Data Jenis Jamur.

Atas perhatian dan kerja sama yang baik diucapkan terima kasih.

Dekan,


 Dr. Eng. Supriatno, ST, MT

Tembusan :

1. Ka BPMPP
2. Mahasiswa
3. File



KELOMPOK TANI NUSA INDAH JATI MAKMUR

Jalan Wijaya Kesuma Gg, Nusa Indah No. 19, Kelurahan Pahlawan,
Kecamatan Binjai Utara, Kota Binjai Telp: 0895-3939-97773

SURAT KETERANGAN NO. 013/SK-KT/08/2024

Binjai, 15 Agustus 2024

Kepada Yth.

Dekan Fakultas Teknik
Universitas Medan Area
Di-
Medan

Dengan Hormat,

Yang bertanda tangan dibawah ini menerangkan bahwa :

Nama : Harya Syafneldi
NPM : 188160105
Program Studi : Teknik Informatika

Adalah benar mahasiswa yang bersangkutan telah selesai melaksanakan penelitian dan pengambilan data di Kelompok Tani Nusa Indah Jati Makmur sebagai syarat penyusunan skripsi yang pelaksanaannya terhitung tanggal 14 Juli s/d 14 Agustus 2024.

Demikian hal ini kami sampaikan agar dapat dipergunakan seperlunya.

Hormat Kami,

Kelompok Tani Nusa Indah Jati Makmur



UNIVERSITAS MEDAN AREA

© Hak Cipta Di Lindungi Undang-Undang

1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area

Document Accepted 26/8/25

Access From (repository.uma.ac.id)26/8/25