

**ANALISIS PERBANDINGAN ALGORITMA *SUPPORT VECTOR MACHINE*,
RANDOM FOREST, DAN *CONVOLUTIONAL NEURAL NETWORK* UNTUK
KLASIFIKASI CITRA PENYAKIT DAUN JAGUNG**



FADHIL HIDAYAT AMIN

H071171311

PROGRAM STUDI SISTEM INFORMASI

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

UNIVERSITAS HASANUDDIN

MAKASSAR



**ANALISIS PERBANDINGAN ALGORITMA *SUPPORT VECTOR MACHINE*,
RANDOM FOREST, DAN *CONVOLUTIONAL NEURAL NETWORK* UNTUK
KLASIFIKASI CITRA PENYAKIT DAUN JAGUNG**

FADHIL HIDAYAT AMIN

H071171311



**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS HASANUDDIN
MAKASSAR
2024**

**ANALISIS PERBANDINGAN ALGORITMA *SUPPORT VECTOR MACHINE*,
RANDOM FOREST, DAN *CONVOLUTIONAL NEURAL NETWORK* UNTUK
KLASIFIKASI CITRA PENYAKIT DAUN JAGUNG**

FADHIL HIDAYAT AMIN

H071171311

Skripsi

Sebagai salah satu syarat untuk mencapai gelar sarjana

Program Studi Sistem Informasi

Pada

29 Juli 2024

**PROGRAM STUDI SISTEM INFORMASI
DEPARTEMEN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGATAHUAN ALAM
UNIVERSITAS HASANUDDIN
MAKASSAR
2024**

SKRIPSI
**ANALISIS PERBANDINGAN ALGORITMA SUPPORT VECTOR MACHINE,
RANDOM FOREST, DAN CONVOLUTIONAL NEURAL NETWORK UNTUK
KLASIFIKASI CITRA PENYAKIT DAUN JAGUNG**

FADHIL HIDAYAT AMIN
H071171311

Skripsi,

Telah dipertahankan di depan Panitia Ujian Sarjana Sistem Informasi pada tanggal 29
Juli 2024 dan dinyatakan telah memenuhi syarat kelulusan

pada

Program Studi Sistem Informasi
Departemen Matematika
Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Hasanuddin
Makassar



Mengesahkan:

Pembimbing Utama,

Dr. Hendra, S.Si., M.Kom.
NIP 19760102 200312 1 001

Pembimbing Pendamping,

A. Muh. Amil Siddik, S.Si., M.Si.
NIP 19911003 201903 1 015

Mengetahui,
Ketua Program Studi,

Prof. Drs. Jeffry Kusuma, Ph.D.
NIP 19641112 198703 1 002

PERNYATAAN KEASLIAN SKRIPSI DAN PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa, skripsi berjudul "Analisis Perbandingan Algoritma *Support Vector Machine*, *Random Forest*, dan *Convolutional Neural Network* Untuk Klasifikasi Citra Penyakit Daun Jagung" adalah benar karya saya dengan arahan dari pembimbing Dr. Hendra, S.Si., M.Kom sebagai Pembimbing Utama dan A. Muh. Amil Siddik, S.Si., M.Si. sebagai Pembimbing Pendamping. Karya ilmiah ini belum diajukan dan tidak sedang diajukan dalam bentuk apa pun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka skripsi ini. Apabila di kemudian hari terbukti atau dapat dibuktikan bahwa sebagian atau keseluruhan skripsi ini adalah karya orang lain, maka saya bersedia menerima sanksi atas perbuatan tersebut berdasarkan aturan yang berlaku.

Dengan ini saya melimpahkan hak cipta (hak ekonomis) dari karya tulis saya berupa skripsi ini kepada Universitas Hasanuddin.

Makassar, 29 Juli 2024



Fadhil Hidayat Amin

H071171311

UCAPAN TERIMA KASIH

Alhamdulillah Rabbil'alamin, Puji dan syukur penulis panjatkan kepada Allah SWT, karena atas berkat, rahmat dan hidayah-Nya sehingga penulis mampu menyelesaikan skripsi dengan judul "**Analisis Perbandingan Algoritma Support Vector Machine, Random Forest, dan Convolutional Neural Network Untuk Klasifikasi Citra Penyakit Daun Jagung**" sebagai syarat untuk menyelesaikan studi Sarjana di Universitas Hasanuddin. Tak lupa pula shalawat serta salam kepada junjungan kita Nabi besar Muhammad Shallallahu Alaihi Wasallam, keluarga, sahabat dan umatnya hingga akhir zaman.

Dalam penyelesaian skripsi ini, penulis menyadari bahwa skripsi ini masih jauh dari kesempurnaan karena adanya hambatan dan ujian yang penulis hadapi selama proses penyelesaian skripsi ini. Oleh karena itu penulis menerima masukan-masukan berupa saran dan kritikan agar karya ilmiah dapat bermanfaat. Selama penyelesaian skripsi ini, penulis banyak menerima doa, bantuan, bimbingan, serta motivasi dari berbagai pihak sehingga segala hambatan yang dihadapi oleh penulis dapat teratasi, oleh karena itu dengan rasa hormat penulis mengucapkan terima kasih yang sebesar-besarnya kepada:

1. Secara istimewa penulis mengucapkan terima kasih untuk orang tua penulis yaitu Ayah **Drs. Raja Amin** dan Ibu **Mutmainnah Mustary, S.Pd., M.Hum.**, yang selama ini telah membesarkan dan memberikan kasih sayang yang tulus, tak henti-hentinya juga mendoakan, menasihati serta mendukung dalam setiap langkah dalam mengejar impian penulis.
2. Kepada adik penulis **Muhammad Fauzan Abrar** yang selama ini selalu mendengar keluh kesah, mendoakan, memotivasi, dan memberikan sumbangan materil kepada penulis sehingga tetap dapat melanjutkan pendidikan hingga sampai ke titik ini.
3. **Prof. Dr. Ir. Jamaluddin Jompa, M.Sc.**, selaku rektor Universitas Hasanuddin yang telah memberikan kesempatan kepada penulis untuk mendapatkan pendidikan.
4. **Dr. Eng. Amiruddin, S.Si., M.Si.**, selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin, beserta para staff yang telah memudahkan penulis dalam mengurus segala hal yang berkaitan dengan persoalan administrasi.
5. **Dr. Firman, S.Si.**, selaku ketua Departemen Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin.
6. **Naimah Aris, S.Si., M.Math.**, selaku Sekretaris Departemen Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin.
7. **Prof. Drs. Jeffry Kusuma, Ph. D.**, selaku Ketua Program Studi Sistem Informasi, Departemen Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin

8. **Dr. Hendra, S.Si., M.Kom.**, selaku pembimbing 1 yang telah banyak memberikan arahan dan nasehat selama penyusunan proposal hingga skripsi.
9. **A. Muh. Amil Siddik, S.Si., M.Si.**, selaku pembimbing 2 yang telah banyak memberikan arahan dan nasehat selama penyusunan proposal hingga skripsi.
10. Terima kasih kepada **Dr. Khaeruddin, M.Sc.** dan **Ir. Eliyah Acantha Manapa Sampetoding, S.Kom., M.Kom.**, selaku dosen penguji yang telah memberikan kritik serta saran Kepada peneliti.
11. **Bapak/Ibu Dosen** Departemen Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin yang telah memberikan saya ilmu dan pengetahuan selama saya belajar di sini.
12. Teman-teman **Sistem Informasi 2017** yang menjadi teman pertama dikampus yang telah memberi dukungan dan kerja samanya dalam menemani saya selama perkuliahan, yang banyak membantu saya dalam hal apapun serta berbagi suka-duka dalam menuntut ilmu. Khususnya, seluruh saudara-saudara **Anak-Anak WICOR, Arya Indrawan Yakub, Restu Adi Akbar, Muh. Ary Putra Ramadhan, Rahmatika, Aris Akhyar Abdillah**, terima kasih atas kebersamaan, loyalitas, pengalaman, dan semangat yang sangat berarti bagi saya.
13. Seluruh informan yang telah berpartisipasi dan meluangkan waktunya untuk membantu penulis dalam proses pengumpulan data sehingga skripsi ini dapat terselesaikan.
14. Terima kasih kepada teman-teman penulis yang telah membantu dalam proses penyusunan skripsi. Mulai dari pengajuan judul, pengumpulan data, proses transkrip hingga tahap penyusunan skripsi sehingga skripsi ini dapat diselesaikan.
15. Kepada seluruh pihak yang tidak dapat saya sebutkan satu persatu, terima kasih sebesar-besarnya untuk segala dukungan dan partisipasi yang diberikan kepada saya semoga bernilai ibadah di sisi Allah Subhanahu Wa Ta'ala.

Penulis mengakui bahwa skripsi ini masih memiliki banyak kekurangan, untuk itu dengan segala kerendahan hati penulis meminta maaf atas kekurangan tersebut. Penutup, semoga tulisan ini dapat memberikan banyak manfaat untuk pembaca. Wassalamu'alaikum Warahmatullahi Wabarakatuh.

Penulis,



Fadhil Hidayat Amin

ABSTRAK

Fadhil Hidayat Amin (H071171311). Analisis Perbandingan Algoritma *Support Vector Machine*, *Random Forest*, dan *Convolutional Neural Network* Untuk Klasifikasi Citra Penyakit Daun Jagung. Dibawah Bimbingan Dr. Hendra, S.Si., M.Kom. dan A. Muh. Amil Siddik, S.Si., M.Si. Program Studi Sistem Informasi, Departemen Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Hasanuddin.

Salah satu tanaman hortikultura yang banyak dibudidayakan oleh petani di Indonesia adalah Jagung (*Zea mays ssp. mays*). Di kecamatan Lau, kabupaten Maros sebagian masyarakat memiliki mata pencaharian sebagai seorang petani tanaman jagung, namun banyak kendala yang telah dihadapi oleh semua petani yakni gagal panen dikarenakan jenis penyakit jagung yang tidak diketahui jenisnya. Penelitian ini bertujuan analisis perbandingan algoritma *Support Vector Machine*, *Random Forest*, dan *Convolutional Neural Network* untuk klasifikasi citra penyakit daun jagung dengan menggunakan beberapa algoritma klasifikasi citra populer yaitu *Support Vector Machine* (SVM), *Random Forest* dan *Convolutional Neural Network* (CNN). Adapun dalam arsitektur uji coba, penulis menggunakan beberapa parameter yang digunakan untuk mengonfigurasi algoritma. Hasil dari uji coba penelitian menunjukkan bahwa CNN memiliki performa terbaik dengan akurasi 96.42%, *precision* (*Blight* 0.96, *Common Rust* 0.99, *Gray Leaf Spot* 0.95), *recall* (*Blight* 0.95, *Common Rust* 1.00, *Gray Leaf Spot* 0.95), dan *F1 Score* (*Blight* 0.96, *Common Rust* 1.00, *Gray Leaf Spot* 0.95).

Kata Kunci: Penyakit Jagung, perbandingan algoritma, *Support Vector Machine* (SVM), *Random Forest*, *Convolutional Neural Network* (CNN), *kaggle*, *google colab*, *Python*.

ABSTRACT

Fadhil Hidayat Amin (H071171311). Comparative Analysis of Support Vector Machine, Random Forest, and Convolutional Neural Network Algorithms for Maize Leaf Disease Image Classification. Under the guidance of Dr. Hendra, S.Si., M.Kom. and A. Muh. Amil Siddik, S.Si., M.Si. Information Systems Study Program, Department of Mathematics, Faculty of Mathematics and Natural Sciences, Hasanuddin University.

One of the horticultural crops widely cultivated by farmers in Indonesia is Maize (*Zea mays* ssp. *mays*). In Lau sub-district, Maros district, some of the community members earn their livelihood as maize farmers, but many obstacles have been faced by all farmers, such as crop failure due to unidentified types of maize diseases. This research aims to conduct a comparative analysis of support vector machine, random forest, and convolutional neural network algorithms for maize leaf disease image classification using several popular image classification algorithms, namely Support Vector Machine (SVM), Random Forest, and Convolutional Neural Network (CNN). In the experimental architecture, the author used several parameters to configure the algorithms. The results of the research experiments showed that CNN has the best performance with an accuracy of 96.42, precision (*Blight* 0.96, *Common Rust* 0.99, *Gray Leaf Spot* 0.95), recall (*Blight* 0.95, *Common Rust* 1.00, *Gray Leaf Spot* 0.95), and F1 Score (*Blight* 0.96, *Common Rust* 1.00, *Gray Leaf Spot* 0.95).

Keywords: Corn Disease, Algorithm Comparison, Support Vector Machine (SVM), Random Forest, Convolutional Neural Network (CNN), Kaggle, Google Colab, Python.

DAFTAR ISI

PERNYATAAN PENGAJUAN	iii
HALAMAN PENGESAHAN	iv
PERNYATAAN KEASLIAN SKRIPSI DAN PELIMPAHAN HAK CIPTA	v
UCAPAN TERIMA KASIH	vi
ABSTRAK	viii
ABSTRACT	ix
DAFTAR ISI	x
DAFTAR TABEL	xii
DAFTAR GAMBAR	xiii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Tujuan Penelitian	3
1.4 Manfaat Penelitian	4
1.5 Batasan Masalah	5
1.6 Tanaman Jagung	5
1.7 Penyakit Pada Tanaman Jagung	6
1.7.1 Blight (Hawar Daun)	6
1.7.2 Common Rust (Karat Daun)	6
1.7.3 Gray Leaf Spot (Bercak Daun)	7
1.8 Bioteknologi	8
1.9 Klasifikasi	9
1.9.1 Support Vector Machine	9
1.9.2 Random Forest	10
1.9.3 Convolutional Neural Network	11
1.10 Google Collab	12
1.11 Kaggle	13
1.12 Evaluasi Kinerja	13
1.12.1 Confusion Matrix	13
1.12.2 Recall	14

1.12.3 Precision	14
1.12.4 F1-Score	15
1.12.5 Accuracy	15
BAB II METODE PENELITIAN.....	16
2.1 Waktu dan Tempat.....	16
2.2 Deskripsi Data	16
2.3 Tahapan Penelitian	16
2.4 Alur Penelitian.....	17
2.5 Instrumen Penelitian.....	18
BAB III HASIL DAN PEMBAHASAN	19
3.1 <i>Dataset</i>	19
3.2 Preprocessing.....	20
3.3 Split Data	21
3.4 Augmentasi Data	22
3.5 Tanaman Jagung	22
3.6 Support Vector Machine	23
3.6.1 Akurasi	24
3.6.2 F1-Score	25
3.6.3 Confusion Matrix	27
3.7 Random Forest.....	29
3.7.1 Akurasi	30
3.7.2 F1-Score	31
3.7.3 Confusion Matrix	32
3.8 Convolutional Neural Network	34
3.8.1 Akurasi	35
3.8.2 F1-Score	36
3.8.3 Confusion Matrix	37
BAB IV PENUTUP	39
4.1 Kesimpulan	39
4.2 Saran	40
DAFTAR PUSTAKA.....	41
LAMPIRAN	47

DAFTAR TABEL

Tabel 1 Confusion Matrix.....	13
Tabel 2 <i>Dataset</i> Corn or Maize Leaf Disease	19
Tabel 3 Pembagian Data.....	21
Tabel 4 Hasil Evaluasi Kinerja SVM.....	28
Tabel 5 Hasil Evaluasi Kinerja Random Forest.....	33
Tabel 6 Hasil Evaluasi Kinerja CNN	38
Tabel 7 Hasil evaluasi kinerja Support Vector Machine	39
Tabel 8 Hasil evaluasi kinerja Random Forest.....	39
Tabel 9 Hasil evaluasi kinerja Convolutional Neural Network.....	40

DAFTAR GAMBAR

Gambar 1 Ilustrasi penyakit <i>Blight</i> (Hawar Daun) pada daun jagung.....	6
Gambar 2 Ilustrasi penyakit <i>Common Rust</i> (Karat Daun) pada daun jagung	6
Gambar 3 Ilustrasi penyakit Grey Leaf Spot (Bercak Daun) pada daun jagung.....	7
Gambar 4 Ilustrasi metode SVM (Fachruddin, 2015).	10
Gambar 5 Ilustrasi Random Forest.....	11
Gambar 6 Ilustrasi Convolutional Neural Network.....	12
Gambar 7 Diagram alur penelitian.	17
Gambar 8 Ilustrasi resize pada citra	20
Gambar 9 Array hasil konversi citra	20
Gambar 10 Array sebelum dan sesudah normalisasi	21
Gambar 11 Proses Augmentasi Data.....	22
Gambar 12 Ilustrasi metode Support Vector Machine (Fachruddin, 2015)	24
Gambar 13 Kurva akurasi Training dan Validation SVM.....	25
Gambar 14 Kurva Training dan Validation F1-Score SVM	26
Gambar 15 Confusion Matrix SVM	28
Gambar 16 Ilustrasi metode Random Forest.....	29
Gambar 17 Kurva akurasi Training dan Validation Random Forest.....	30
Gambar 18 Kurva F1-Score Training dan Validation Random Forest	31
Gambar 19 Confusion Matrix Random Forest	33
Gambar 20 Ilustrasi metode Convolutional Neural Network.....	34
Gambar 21 Kurva akurasi Training dan Validation CNN.....	35
Gambar 22 Kurva F1-Score Training dan Validation CNN	36
Gambar 23 Confusion Matrix CNN	38

BAB I

PENDAHULUAN

1.1 Latar Belakang

Indonesia merupakan negara agraris yang mayoritas penduduknya bermata pencaharian sebagai petani. Hal ini ditunjang dari banyaknya lahan kosong yang dapat dimanfaatkan sebagai lahan pertanian. Sektor pertanian merupakan sektor yang mempunyai peranan strategis dan struktur pembangunan bagi perekonomian nasional. Sektor pertanian di Indonesia meliputi subsektor hortikultura, subsektor perikanan, subsektor peternakan, dan subsektor kehutanan. Salah satu subsektor pertanian yang mempunyai nilai ekonomi tinggi dan memiliki potensi untuk terus dikembangkan adalah subsektor hortikultura (Hortikultura, 2020). Secara harfiah istilah Hortikultura diartikan sebagai usaha membudidayakan tanaman buah-buahan, sayuran dan tanaman hias.

Salah satu tanaman hortikultura yang banyak dibudidayakan oleh petani di Indonesia adalah Jagung (*Zea mays ssp. mays*). Di kecamatan Lau, kabupaten Maros sebagian masyarakat memiliki mata pencaharian sebagai seorang petani tanaman jagung, namun banyak kendala yang telah dihadapi oleh semua petani yakni gagal panen dikarenakan jenis penyakit yang tidak diketahui jenisnya. Banyaknya penyakit jagung yang tidak diketahui jenisnya menyebabkan banyak kerugian bagi para petani.

Penyakit jagung dapat dikenali dari penampilan daunnya. Adapun daun yang berpenyakit akan mempunyai ciri-ciri dari segi warna dan tekstur daun. Beragamnya penyakit daun jagung mempunyai gejala fisik misalnya warna daun kuning, daun bercak-bercak, perubahan tekstur daun dan perubahan bentuk daun. Dari gejala tersebut dapat di klasifikasi menjadi Bercak Daun (*Bipolaris maydis* Syn), Karat Daun (*Puccinia polysora*), dan Hawar Daun (*Rhizoctonia solani*) (Sudjono, 1988).

Terdapat beberapa algoritma klasifikasi citra populer yaitu *Support Vector Machine* (SVM), *Random Forest* dan *Convolutional Neural Network* (CNN). *Support Vector Machine* dan *Random Forest* merupakan algoritma klasifikasi dari *Machine Learning* sedangkan *Convolutional Neural Network* merupakan algoritma klasifikasi dari *Deep Neural Network*. Pada *Machine Learning* data akan di training dan testing sedangkan pada *Deep Learning* akan menggunakan banyak lapisan ketika mengolah data.

Ada beberapa penelitian sebelumnya yang telah melakukan klasifikasi menggunakan algoritma terkait CNN, SVM, dan *Random Forest*. Penelitian pertama yakni Klasifikasi Penyakit Daun Jagung Menggunakan Metode *Convolutional Neural Network* pada April 2022 yang ditulis oleh Ivan Pratama Putra, Rusbandi, dan Derry Alamsyah (Putra, Rusbandi, & Alamsyah, 2022). Penelitian tersebut memiliki kesimpulan yakni hasil penelitian yang didapat tentang klasifikasi penyakit daun jagung menggunakan metode *Convolutional Neural Network* (CNN) yang telah dilakukan, dengan menggunakan *model Resnet 50* dan *Optimizer Adam* mendapatkan akurasi yang paling tinggi sebesar 98,4% daripada Optimizer lainnya.

Penelitian berikutnya yaitu berjudul Klasifikasi Penyakit Daun Padi menggunakan *Random Forest* dan *Color Histogram* yang ditulis oleh Sarifah Agustiani, Yoseph Tajul Arifin, Agus Junaidi, Siti Khotimatul Wildah dan Ali Mustopa pada Juni 2022 (Agustiani, et al., 2022). Penelitian yang dirancang oleh Sarifah dan kawan-kawan memiliki kesimpulan bahwa dari beberapa skenario pengujian yang dilakukan mulai dari perbandingan dengan beberapa ekstraksi fitur yang berbeda, perbandingan dengan beberapa algoritma yang beda, perbandingan antara menggunakan teknik augmentasi maupun tidak. Hasil akurasi tertinggi diperoleh dari metode yang diusulkan yaitu penggunaan ekstraksi fitur *Color Histogram* serta algoritma *Random Forest* yang mampu menghasilkan nilai akurasi sebesar 99.65%. Dengan nilai akurasi yang dihasilkan semakin tinggi akan menunjukkan kemampuan pengklasifikasian yang semakin akurat.

Kemudian penelitian berikutnya yaitu berjudul Klasifikasi Jenis Pisang Berdasarkan Fitur Warna, Tekstur, Bentuk Citra Menggunakan SVM dan KNN yang ditulis oleh Yusuf Eka Yana dan Nur Nafi'iyah pada April 2021 (Yana & Nafi'iyah, 2021). Penelitian yang dirancang oleh Yusuf dan Nur memiliki kesimpulan hasil klasifikasi jenis Pisang dengan algoritma *Support Vector Machine* yang paling baik adalah fitur warna 41,67%. Hasil klasifikasi jenis Pisang dengan algoritma KNN, K=2 nilai fitur tekstur paling baik hasil akurasinya 58,33%.

Dari beberapa penelitian di atas, penulis berencana menyusun tugas akhir dengan judul "ANALISIS PERBANDINGAN ALGORITMA *SUPPORT VECTOR MACHINE*, *RANDOM FOREST*, DAN *CONVOLUTIONAL NEURAL NETWORK* UNTUK KLASIFIKASI PENYAKIT DAUN JAGUNG". Meskipun algoritma *Machine Learning* telah banyak diterapkan pada klasifikasi gambar, suatu algoritma tidak selalu cocok untuk setiap tipe data atau tipe gambar. Oleh karena itu penelitian ini bertujuan untuk membandingkan algoritma mana yang lebih baik untuk mendeteksi klasifikasi penyakit daun jagung antara *Support Vector Machine*, *Random Forest* dan *Convolutional Neural Network*.

1.2 Rumusan Masalah

Berdasarkan latar belakang di atas, dapat dirumuskan beberapa masalah sebagai berikut:

- 1) Bagaimana performa *Support Vector Machine* dalam mengelompokkan citra penyakit daun jagung?
- 2) Bagaimana performa *Random Forest* dalam mengelompokkan citra penyakit daun jagung?
- 3) Bagaimana performa *Convolutional Neural Network* dalam mengelompokkan citra penyakit daun jagung?
- 4) Bagaimana mendapatkan performa terbaik dari hasil perbandingan performa metode *Support Vector Machine*, *Random Forest*, dan *Convolutional Neural Network* dalam mengelompokkan citra penyakit daun jagung?

1.3 Tujuan Penelitian

Dengan memperhatikan latar belakang dan rumusan masalah di atas, maka tujuan penelitian ini adalah:

- 1) Untuk mengetahui bagaimana performa *Support Vector Machine* dalam mengelompokkan citra penyakit daun jagung.
- 2) Untuk mengetahui bagaimana performa *Random Forest* dalam mengelompokkan citra penyakit daun jagung.
- 3) Untuk mengetahui bagaimana performa *Convolutional Neural Network* dalam mengelompokkan citra penyakit daun jagung.
- 4) Untuk mengetahui bagaimana mendapatkan performa terbaik dari hasil perbandingan performa metode *Support Vector Machine*, *Random Forest*, dan *Convolutional Neural Network* dalam mengelompokkan citra penyakit daun jagung.

1.4 Manfaat Penelitian

Manfaat penelitian dalam penulisan proposal ini adalah:

- 1) Dapat mengetahui metode yang terbaik di antara metode *Support Vector Machine*, *Random Forest*, dan *Convolutional Neural Network* dalam mengelompokkan citra penyakit daun jagung.
- 2) Dapat digunakan sebagai model prediksi dasar untuk penelitian masa depan.

Manfaat penelitian untuk masyarakat adalah:

- 1) Dapat digunakan sebagai sumber pangan. Jagung merupakan salah satu sumber pangan bagi banyak masyarakat di seluruh dunia, terutama di negara-negara berkembang. Berbagai produk pangan seperti tepung jagung, *popcorn* dan mi jagung menjadi bagian penting dari pola makan sehari-hari.
- 2) Dapat digunakan sebagai sumber kesehatan. Jagung mengandung serat, vitamin, dan mineral yang penting untuk kesehatan tubuh manusia. Konsumsi jagung dapat membantu menjaga kesehatan pencernaan, mengurangi risiko penyakit jantung, serta mendukung kesehatan mata dan kulit.
- 3) Dapat digunakan sebagai sumber pendapatan petani. Jagung menjadi sumber pendapatan utama bagi banyak petani di berbagai negara. Pertanian jagung memberikan lapangan kerja dan meningkatkan perekonomian lokal di banyak daerah.

Manfaat penelitian bagi teknologi adalah:

- 1) Dapat digunakan sebagai bioteknologi. Perkembangan dalam bidang bioteknologi telah memungkinkan pengembangan varietas jagung yang lebih tahan terhadap hama dan penyakit, serta memiliki hasil yang lebih tinggi. Hal ini membantu meningkatkan produktivitas pertanian dan ketahanan pangan.
- 2) Dapat digunakan sebagai industri. Jagung juga digunakan dalam berbagai industri, termasuk industri makanan, minuman, farmasi, dan kosmetik. Penggunaan jagung sebagai bahan baku dalam industri membantu menciptakan lapangan kerja dan meningkatkan pertumbuhan ekonomi.

1.5 Batasan Masalah

Adapun batasan masalah yang diteliti untuk mencegah pembahasan yang terlalu luas, yaitu:

- 1) *Dataset* penyakit daun jagung berasal dari web resmi *Kaggle (Kaggle.com)* yaitu *corn or maize leaf disease dataset*.
- 2) Penyakit yang diteliti ialah penyakit pada tanaman daun jagung di antaranya *Blight* (Hawar Daun), *Common Rust* (Karat Daun), dan *Grey Leaf Spot* (Bercak Daun).

1.6 Tanaman Jagung

Di dunia ini jagung (*Zea mays ssp. mays*) adalah tanaman pangan ketiga setelah gandum dan padi. Jagung berasal dari Mexico, disana telah dibudidayakan selama ribuan tahun. Jagung menjadi dasar dari kebudayaan *Aztec* dan *Maya*. Amerika Serikat merupakan produsen jagung terbesar karena menghasilkan lebih dari setengah produksi dunia (Syarifudin, Hidayat, & Fanani, 2017).

Di Indonesia jagung adalah tanaman pangan kedua setelah padi. Tanaman jagung merupakan salah satu jenis tanaman pangan biji-bijian dari keluarga rumput-rumputan yang berasal dari Amerika yang tersebar ke Asia dan Afrika melalui kegiatan bisnis orang-orang Eropa ke Amerika (Khair, Pasaribu, & Suprpto, 2013). Jagung dapat disebut juga sebagai salah satu tanaman serealia penting di Indonesia, selain sebagai tanaman bahan pangan pokok pengganti beras dalam upaya diversifikasi pangan, jagung juga merupakan pakan ternak (Kurniawati, Yulianingsih, & Wahda, 2021).

Jagung memiliki banyak manfaat bagi tubuh karena kandungan nutrisinya, di antara manfaatnya yaitu dapat menurunkan *hipertensi* sehingga dapat mencegah penyakit jantung. Jagung dapat mengontrol *diabetes*, memperlancar pencernaan, mencegah sembelit dan wasir karena jagung kaya akan serat, bahkan dapat menurunkan risiko kanker usus besar. Jagung juga mengandung sebagian besar *magnesium*, zat besi, dan yang terpenting adalah kandungan *fosfor* yang baik untuk kesehatan tulang, nutrisi ini tidak hanya mencegah tulang retak karena pertambahan usia, tapi juga meningkatkan fungsi ginjal. *Vitamin C*, *karotenoid* dan *bioflavinoids* yang terkandung dalam jagung juga dapat menjaga jantung agar tetap sehat dengan mengendalikan kadar kolesterol dan meningkatkan aliran darah dalam tubuh (Fikrie, 2019).

1.7 Penyakit Pada Tanaman Jagung

1.7.1 *Blight* (Hawar Daun)



Gambar 1 Ilustrasi penyakit *Blight* (Hawar Daun) pada daun jagung

Berdasarkan ilustrasi pada gambar 1, awal terinfeksi hawar daun menunjukkan gejala berupa bercak kecil, berbentuk oval. Infeksinya berkembang sehingga bercak semakin memanjang berbentuk *ellips* dan berkembang menjadi *nekrotik* (disebut hawar), warnanya hijau keabu-abuan atau coklat (Wakman & Burhanuddin, 2007). Infeksi berat akibat serangan penyakit hawar daun dapat mengakibatkan tanaman jagung cepat mati atau mengering. Hawar daun disebabkan oleh jamur *Helminthosporium turcicum* dan akan terus hidup selama inang masih mendapat suplai makanan. Panjang hawar 2,5-15 cm, bercak muncul di mulai dari daun terbawah kemudian berkembang menuju daun atas. Jamur ini tidak menginfeksi tongkol atau klobot jagung (Syarifudin, Hidayat, & Fanani, 2017).

1.7.2 *Common Rust* (Karat Daun)

Penyakit *Common Rust* atau karat daun pada jagung ini di Indonesia baru menarik perhatian pada tahun 1950-an. Penyakit karat daun sudah terdapat pada bahan yang dikumpulkan Van Der Goot di Bogor pada tahun 1923 dan yang dikumpulkan oleh Schwarz dari Bandung pada tahun 1925. Penyakit karat dapat terjadi di dataran rendah sampai tinggi, infeksi berkembang baik pada musim penghujan atau musim kemarau (Syarifudin, Hidayat, & Fanani, 2017).



Gambar 2 Ilustrasi penyakit *Common Rust* (Karat Daun) pada daun jagung

Berdasarkan ilustrasi pada gambar 2, gejala pada karat daun ini berupa bercak-bercak kecil (*uredinia*) berbentuk bulat sampai oval terdapat di permukaan daun jagung bagian atas maupun bawah. *Uredinia* menghasilkan *uredospora* berbentuk bulat atau oval serta berperan penting sebagai sumber *inokulum* dalam menginfeksi tanaman jagung lainnya (Syarifudin, Hidayat, & Fanani, 2017). Penyakit ini dapat menyebar dengan cepat diketahui melalui perantara angin, sehingga menyebabkan petani harus lebih waspada terhadap penyakit ini. Penyakit karat daun disebabkan oleh jamur *Puccinia sorghi*. Gejala awal berupa bercak-bercak merah dan keluar serbuk seperti tepung berwarna coklat kekuningan. Akibat penyakit ini, tanaman tidak dapat melakukan fotosintesis dengan sempurna sehingga pertumbuhannya melambat, bahkan tanaman dapat mati (Prasetyo, Ratih, Ivayani, & Akin, 2017).

1.7.3 Gray Leaf Spot (Bercak Daun)



Gambar 3 Ilustrasi penyakit *Grey Leaf Spot* (Bercak Daun) pada daun jagung

Berdasarkan ilustrasi pada gambar 3, gejala visual penyakit bercak daun menurut (Wakman & Burhanuddin, 2007) adalah adanya lesio pada daun jagung biasanya memanjang diantara tulang daun dengan warna coklat muda. Lesio biasanya muncul dengan ukuran mencapai 1,2 x 2,7 cm, berbentuk *ellips*, dan lesio sering dikelilingi oleh warna coklat dan dapat juga muncul pada bagian batang (Talanca & Tenrirawe, 2015). Penyakit *Grey Leaf Spot* atau biasa disebut Bercak daun, umumnya berwarna hijau atau coklat pada bagian atas daun dan dapat menjalan ke bagian bawah menuju batang sehingga menyebabkan tanaman menjadi kering dan mati.

Penyakit ini disebabkan oleh patogen *Bipolaris maydis* Syn. Penyakit bercak daun mulai muncul setelah tanaman jagung berumur sekitar 2 minggu setelah tanam. Tanaman jagung yang terserang menjadi layu atau mati dalam waktu 3-4 minggu. Penyakit bercak daun sudah menyebar di seluruh sentra penghasil jagung di Indonesia seperti di Jawa, Kalimantan, Sumatera, dan Sulawesi. Dalam (Pakki, 2005) kehilangan hasil akibat serangan penyakit bercak daun pada tanaman jagung dapat mencapai 59%.

1.8 Bioteknologi

Bioteknologi adalah bidang ilmu yang mempelajari tentang makhluk hidup, baik *mikroorganisme* maupun *makroorganisme*, serta pemanfaatannya dalam berbagai aspek kehidupan. Bioteknologi mencakup penerapan praktis organisme hidup atau komponen subselulernya dalam industri jasa dan manufaktur, serta pengelolaan lingkungan. Bioteknologi dibagi menjadi bioteknologi konvensional dan *modern*. Bioteknologi konvensional menggunakan *mikrobia* untuk memodifikasi bahan dan lingkungan untuk memperoleh produk optimal, seperti pembuatan tempe, tape, roti, dan pengomposan sampah. Sementara itu, bioteknologi *modern* menggunakan ketrampilan manusia dalam melakukan manipulasi makhluk hidup untuk menghasilkan produk sesuai yang diinginkan manusia, seperti melalui rekombinan, yaitu pembentukan kombinasi materi genetik yang baru dengan cara penyisipan molekul DNA ke dalam suatu vektor (Prof. Drs. Sutarno, 2016).

Definisi lain menerangkan bahwa Bioteknologi adalah metode atau proses yang melibatkan makhluk hidup atau organisme hidup untuk menghasilkan produk baru sehingga dapat bermanfaat bagi kesejahteraan manusia. Atau dapat dikatakan juga, bioteknologi merupakan suatu cabang ilmu yang mempelajari cara memanfaatkan organisme hidup dalam melakukan proses produksi untuk menghasilkan barang maupun jasa yang bermanfaat bagi manusia. Bioteknologi berasal dari dua kata yaitu Bio dan Teknologi. Kata Bio berarti kehidupan sedangkan kata Teknologi memiliki makna sebagai suatu metode ilmiah yang digunakan untuk mencapai tujuan secara praktis. Pada zaman kini, Bioteknologi tidak hanya didukung oleh keilmuan biologi saja, tetapi juga dari berbagai macam ilmu terapan, seperti biokimia, biologi molekuler, genetika, mikrobiologi, komputer dan lain-lain. Sehingga bioteknologi dapat didefinisikan sebagai ilmu terapan yang menggabungkan berbagai macam cabang ilmu dalam memproses barang atau jasa yang bisa bermanfaat bagi manusia dengan menggunakan bantuan makhluk hidup. Adapun ciri-ciri utama dari bioteknologi adalah adanya (Amar, Nadia, & Sartika, 2020):

- 1) Jasad biologi/ renik seperti mikroorganisme, tumbuhan maupun hewan.
- 2) Teknologi yang dapat diaplikasikan secara industri, baik dengan menggunakan proses ekstraksi maupun pemurnian.

Bioteknologi juga digunakan dalam berbagai bidang ilmu, seperti teknik lingkungan, teknik pangan, teknik kedokteran, dan lain sebagainya. Dalam bidang pangan, bioteknologi digunakan untuk menghasilkan produk olahan pangan susu (keju) yang halal, menggunakan *mikroorganisme* seperti *Rhizopus oryzae*. Bioteknologi modern juga menghasilkan produk seperti *enzim*, *glukosa* hasil *hidrolisis enzimatis*, dan beberapa bahan tambahan pangan serta produk hasil rekayasa genetika (Estikomah, 2020).

Bidang IT (*Information Technology*) juga memiliki peran penting dalam pengembangan bioteknologi. IT digunakan dalam pengembangan alat-alat dan metode baru yang memanfaatkan teknologi bioteknologi. Dalam pengajaran bioteknologi, IT membantu dalam meningkatkan kualitas pembelajaran dengan menggunakan berbagai teknologi, seperti multimedia dan aplikasi pembelajaran interaktif. Hal ini membantu siswa dalam memahami konsep bioteknologi dengan lebih baik dan meningkatkan keterampilan mereka dalam bidang ini (Wahyuni, Noviani, Sartika, Habibie, & Handayani, 2024).

1.9 Klasifikasi

Klasifikasi merupakan bagian dari metode prediksi, namun yang diprediksi adalah *label* atau *target* dari sebuah *tuple data*. Dengan kata lain, tujuan klasifikasi adalah menerima suatu input sejenis dengan *dataset* awal, kemudian menentukan kelas *target* atau *output*-nya (Kotsiantis, Zaharakis, & Pintelas, 2007). Menurut Goronescu (2011), dalam proses klasifikasi didasarkan oleh empat komponen dasar yang sangat penting, yaitu (Goronescu, 2011):

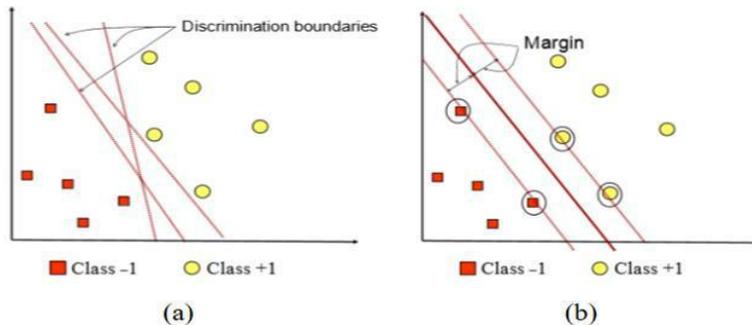
1. *Training dataset*, yakni data yang digunakan untuk melatih model dalam mengenali kelas-kelas berdasarkan variabel independen dan variabel dependen yang ada.
2. *Validation dataset*, yaitu data yang digunakan untuk menguji validasi model *classifier* yang telah dilatih pada proses sebelumnya.

Target class merupakan variabel dependen yang dipengaruhi oleh *predictor* pada suatu *dataset*.

1.9.1 Support Vector Machine

Support Vector Machine (SVM) merupakan metode untuk menyelesaikan masalah statistika klasik, khususnya pada *task* klasifikasi dan prediksi. SVM sendiri dikembangkan menggunakan prinsip *linear classifier*. Namun seiring berkembangnya metodologi klasifikasi, sering pula dijumpai data yang bersifat *non-linear*, sehingga dikembangkan SVM dengan konsep *kernel* untuk pemetaan yang akurat (Hsu, Chang, & Lin, 2003).

Support Vector Machine (SVM) adalah sebuah algoritma pembelajaran mesin yang digunakan untuk klasifikasi dan regresi. SVM bekerja berdasarkan konsep Structural Risk Minimization (SRM) yang bertujuan untuk mengolah data menjadi hyperplane yang dapat memisahkan ruang input menjadi dua kelas. Dalam prosesnya, SVM mencari hyperplane terbaik yang dapat memisahkan titik data dengan jarak maksimum. Hal ini membuat SVM sangat efektif dalam mengatasi masalah klasifikasi dan regresi, baik dalam bentuk linear maupun non-linear (Huang & Putra, 2024).



Gambar 4 Ilustrasi metode SVM (Fachruddin, 2015).

Konsep SVM adalah menemukan *hyperplane* pada ruang *input*, dan *hyperplane*-nya sendiri digunakan sebagai pemisah antara dua kelas pada ruang *input*. Pada gambar 4(a), tiga garis merah disebut sebagai *discriminant boundaries* atau garis pemisah yang dihasilkan untuk kelas yang ada, dan *pattern* yang paling dekat disebut sebagai *support vector*. Sedangkan garis tebal pada gambar 4(b) menunjukkan *hyperplane* terbaik antar kelas, karena terletak tepat di antara kedua kelas. (Fachruddin, 2015)

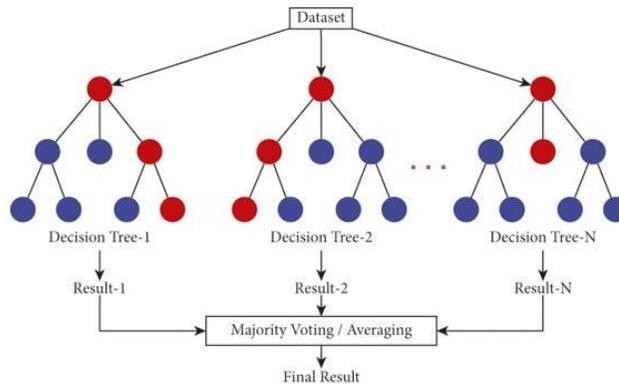
SVM juga memiliki keuntungan lain, seperti kemampuan untuk menangani data berdimensi tinggi dan tidak mengalami kondisi *overfitting*. Algoritma ini juga dapat digunakan pada model pembelajaran semi-supervised dan memiliki kompleksitas komputasi yang lebih baik dibandingkan dengan algoritma lain. Selain itu, SVM dapat menangani data yang tidak berdistribusi teratur dan tidak diketahui distribusinya, membuatnya sangat berguna dalam situasi yang kompleks (Team, 2021).

Dalam kesimpulan, SVM merupakan algoritma pembelajaran mesin yang sangat efektif dalam mengatasi masalah klasifikasi dan regresi. Dengan kemampuan menangani data berdimensi tinggi, tidak mengalami *overfitting*, dan dapat digunakan pada model pembelajaran semi-supervised, SVM menjadi pilihan yang tepat untuk berbagai aplikasi. Sumber jurnal yang mendukung keefektifan SVM antara lain penelitian pengelasan logam dan klasifikasi akreditasi sekolah dasar (Octaviani, Wilandari, & Ispriyanti, 2014).

1.9.2 Random Forest

Random Forest (RF) adalah kombinasi *predictor* berbentuk *tree*, dan tiap *tree*-nya bergantung pada nilai suatu *vektor* acak berdasarkan *sampling* secara independen dengan distribusi yang setara (Breiman, 2001). Proses *randomization* untuk membentuk *tree* tidak hanya dilakukan pada data sampel saja, melainkan juga pada pengambilan variabel *predictor*-nya, hal ini akan menghasilkan kumpulan *classification tree* dengan ukuran dan bentuk yang berbeda-beda. Salah satu dari kelebihan *Random Forest* adalah dapat digunakan baik untuk

menyelesaikan masalah klasifikasi maupun masalah regresi. Masalah tersebut merupakan sistem *machine learning* saat ini. Selain itu, *Random Forest* dikenal dapat digunakan pada *data imbalance* di *class* yang berbeda, khususnya untuk *dataset* yang berukuran besar (Paul, et al., 2018).



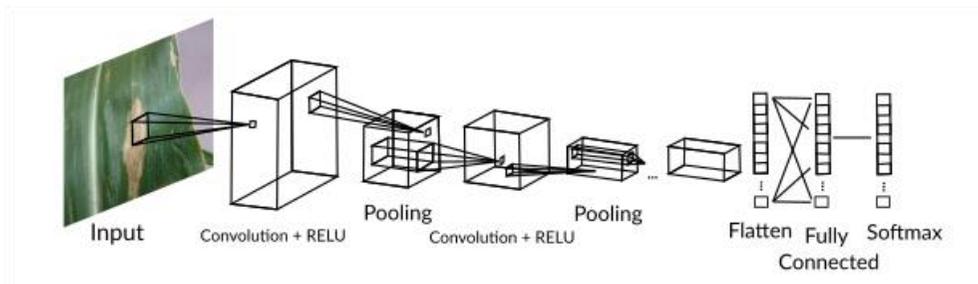
Gambar 5 Ilustrasi *Random Forest*

Berdasarkan ilustrasi pada gambar 5, dapat dilihat bahwa RF adalah *classifier* yang terdiri dari beberapa *decision tree*, yang masing-masing berperan pada *class* tertentu. Dengan menggabungkan beberapa *tree* di RF dapat menghasilkan prediksi yang lebih baik (Devetyarov & Nourtdinov, 2010). Hasil yang diharapkan adalah suatu kumpulan *classification tree* yang memiliki korelasi kecil antar *tree*-nya. Korelasi yang kecil akan menurunkan hasil kesalahan prediksi *Random Forest* (Breiman, 2001).

Salah satu kelebihan *Random Forest* adalah kemampuannya untuk menangani fitur-fitur kategori. Dalam pengolahan data, fitur-fitur kategori sering kali memerlukan encoding khusus untuk dapat digunakan dalam model *machine learning*. Namun, *Random Forest* dapat menangani fitur-fitur kategori secara langsung tanpa perlu encoding, sehingga memudahkan proses pengolahan data (Deng, 2018).

1.9.3 Convolutional Neural Network

Convolutional Neural Network merupakan salah satu metode dalam *Deep Learning*. Metode ini berasal dari pengembangan dari *Multilayer Perceptron* (MLP) tradisional dan biasa digunakan untuk mengolah data dengan bentuk citra. Salah satu metode dalam pengklasifikasian citra (*images*) adalah dengan menggunakan *Convolutional Neural Networks* (CNN) (Putra, Rusbandi, & Alamsyah, 2022).



Gambar 6 Ilustrasi *Convolutional Neural Network*

CNN merupakan kombinasi dari jaringan syaraf tiruan dan metode *Deep learning* (Xiaofeng & Yan, 2015). CNN terdiri dari satu atau lebih lapisan konvolusal, seringkali dengan suatu lapisan subsampling yang diikuti oleh satu atau lebih lapisan yang terhubung penuh sebagai standar jaringan syaraf (Hijazi, Kumar, & Rowen, 2015). Metode CNN sangat cocok dalam proses ekstraksi fitur yang terdapat dalam suatu citra melalui lapisan konvolusi yang merupakan ciri khas dari metode ini, sehingga dapat terbentuk model yang dapat mengenali fitur-fitur yang sangat kompleks dengan detail. Tetapi dalam proses pembuatan model ini dibutuhkan sumber daya yang cukup besar agar dapat melakukan proses komputasi dalam pembuatan model, serta memakan waktu yang cukup lama tergantung dalam jumlah data, ukuran, dan warna citra yang digunakan. Maka proses pembuatan model dengan metode CNN ini biasanya menggunakan GPU agar dapat melakukan proses pembuatan model lebih cepat.

Convolutional Neural Networks (CNN) memiliki beberapa manfaat dalam bidang ilmu Sistem Informasi. Contoh Pertama yakni untuk ekstraksi Fitur, CNN memiliki ekstraksi fitur yang kuat, yang memungkinkannya untuk mengolah dan mengklasifikasikan citra dengan baik. Hal ini sangat berguna dalam bidang Sistem Informasi untuk mengolah dan menganalisis data citra. Kedua yaitu untuk Pengenalan Citra, CNN dirancang khusus untuk pengenalan dan klasifikasi gambar. Aplikasi ini menggunakan bahasa pemrograman *Python*, *TensorFlow*, dan *OpenCV*. Proses *training* dilakukan dengan menggunakan *batch size* dan *epoch* yang berbeda-beda (Darmawan, 2023).

1.10 Google Collab

Google Colab environment dari *Google* yang digunakan untuk menjalankan program. *Colab* adalah salah satu produk dari *Google* yang berbasis *cloud* dan dijalankan melalui *browser*. *Colab* menyediakan *processor* dengan spesifikasi tinggi (GPU dan TPU) dengan tujuan memudahkan para *researcher* menjalankan program yang membutuhkan spesifikasi tinggi secara online (Bisong, 2019). *Google* menyediakan GPU (graphics processing unit) yang bisa digunakan secara gratis dengan tipe Nvidia K80s, P4s, T4s, dan untuk versi berbayar akan mendapatkan tipe V100 dan P100s (Setiawan, 2021).

1.11 Kaggle

Kaggle merupakan salah satu *website* yang dapat digunakan untuk belajar *data science* dengan mudah. Situs ini memberikan berbagai jenis fitur menarik dalam implementasi *data science*. *Kaggle* merupakan salah satu situs penyedia *dataset* dan kode-kode *data science* yang dapat kita pelajari kapanpun dan dimanapun (Mangkunegara & Purwono, 2022).

Kaggle dilengkapi dengan beberapa menu utama yaitu:

- Competitions* dikhususkan bagi anda yang suka melakukan kompetisi pada kasus *data science*.
- Datasets* merupakan menu *kaggle* yang menyajikan berbagai jenis *dataset* untuk *data science*.
- Code* merupakan sejenis *notebook* ataupun komputer pribadi untuk menjalankan kode-kode *data science*.
- Discussions* merupakan wadah komunitas pengguna *kaggle* untuk berdiskusi terkait berbagai jenis permasalahan *data science*.
- Courses* merupakan berbagai jenis kursus *online* yang tersedia di *kaggle.com*.

Situs *kaggle* dapat diakses pada laman *kaggle.com*. *Kaggle* menyediakan *dataset* dan *notebook* dalam satu paket komputasi yang dapat dijalankan secara *online*. Banyak *programmer data science* atau seorang *data scientist* memanfaatkan situs ini untuk mengolah data mereka.

1.12 Evaluasi Kinerja

1.12.1 Confusion Matrix

Confusion matrix merupakan pengukuran performa yang digunakan pada masalah klasifikasi, *confusion matrix* dapat berupa keluaran visualisasi yang menunjukkan hasil prediksi dari model yang benar maupun salah (Nuel Jaya, 2022).

Tabel 1 Confusion Matrix

		Prediksi		
		Kelas 1	Kelas 2	Kelas 3
Aktual	Kelas 1	TP(T ₁₁)	FN(F ₁₂)	FN(F ₁₃)
	Kelas 2	FP(F ₂₁)	TP(T ₂₂)	FN(F ₂₃)
	Kelas 3	FP(F ₃₁)	FP(F ₃₂)	TP(T ₃₃)

Dalam tabel di atas :

- TP(T₁₁), TP(T₂₂), dan TP(T₃₃) mewakili *True Positives* untuk kelas 1, 2, dan 3 masing-masing.
- FN(F₁₂) mewakili *False Negatives* untuk kelas 1 (diprediksi sebagai kelas 2).
- FP(F₂₁) mewakili *False Positives* untuk kelas 2 (diprediksi sebagai kelas 1).
- FN(F₁₃) mewakili *False Negatives* untuk kelas 1 (diprediksi sebagai kelas 3).
- FP(F₃₁) mewakili *False Positives* untuk kelas 3 (diprediksi sebagai kelas 1).
- FN(F₂₃) mewakili *False Negatives* untuk kelas 2 (diprediksi sebagai kelas 3).
- FP(F₃₂) mewakili *False Positives* untuk kelas 3 (diprediksi sebagai kelas 2).

1.12.2 Recall

Recall merupakan rasio prediksi benar positif (*True Positives*) dibandingkan dengan keseluruhan data yang benar positif (*True Positives*). Berikut rumus *recall* yang digunakan (Nuel Jaya, 2022).

$$Recall (i) = \frac{T_i}{T_i + \sum_{i=1}^C F_{ij}}$$

Keterangan:

- T(i) = Jumlah *True Positive* untuk kelas (i)
- F(ij) = Jumlah *False Negative* untuk kelas (i) yang diprediksi sebagai kelas (j) (j ≠ i)
- C = Jumlah total kelas

1.12.3 Precision

Precision merupakan rasio prediksi benar positif (*True Positives*) dibandingkan dengan keseluruhan hasil yang diprediksi positif (*True Positives*) (Nuel Jaya, 2022).

$$Precision (i) = \frac{T_i}{T_i + \sum_{j=1}^C F_{ji}}$$

Keterangan :

- T(i) = Jumlah *True Positive* untuk kelas (i)
- F(ji) = Jumlah *False Positive* untuk kelas (j) yang diprediksi sebagai kelas (i) (j ≠ i)
- C = Jumlah total kelas

1.12.4 F1-Score

F1-Score merupakan perbandingan rata-rata *precision* dan *recall*. Berikut rumus *F1-Score* yang digunakan (Nuel Jaya, 2022).

$$Accuracy = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

1.12.5 Accuracy

Accuracy merupakan rasio prediksi benar (positif dan negatif) dengan keseluruhan data. Biasanya akurasi suatu model ditentukan dalam bentuk persen (Nuel Jaya, 2022).

$$Accuracy = \frac{\sum_{i=1}^C T_i}{N}$$

Keterangan :

- T(i) = Jumlah True Positive untuk kelas (i)
- N = Jumlah total sampel
- C = Jumlah total kelas

BAB II

METODE PENELITIAN

2.1 Waktu dan Tempat

Penelitian ini dilaksanakan terhitung dari waktu untuk melaksanakan perencanaan penelitian, proses pelaksanaan penelitian, sampai dengan akhir proses pembuatan laporan lengkap mengenai penelitian. Lokasi penelitian dilakukan di Laboratorium Rekayasa Perangkat Lunak Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Hasanuddin.

2.2 Deskripsi Data

Data diambil dari *website* resmi *Kaggle* (*kaggle.com*). Data yang digunakan merupakan *dataset* citra dari *kaggle* yakni *Corn or Maize Leaf Disease Dataset* (<https://www.kaggle.com/datasets/smaranjitghose/corn-or-maize-leaf-disease-dataset>). Data ini berisi empat kelas gambar yaitu *Grey Leaf Spot* (Bercak Daun) berjumlah 1146 *images*, *Common Rust* (Karat Daun) berjumlah 1306 *images*, *Blight* (Hawar Daun) berjumlah 574 *images*, dan *Healthy* (Daun Sehat) berjumlah 1162 *images*. Data yang digunakan yaitu *Grey Leaf Spot* (Bercak Daun) berjumlah 574 *images*, *Common Rust* (Karat Daun) berjumlah 574 *images*, *Blight* (Hawar Daun) berjumlah 574 *images*. Sehingga total jumlah data yang digunakan berjumlah 1722 *images* dengan format *jpg*.

2.3 Tahapan Penelitian

Untuk menyelesaikan penelitian ini, peneliti akan melewati beberapa tahap penelitian, yaitu: Pra-penelitian, eksplorasi data, *set data Training*, *Validation*, dan analisis hasil.

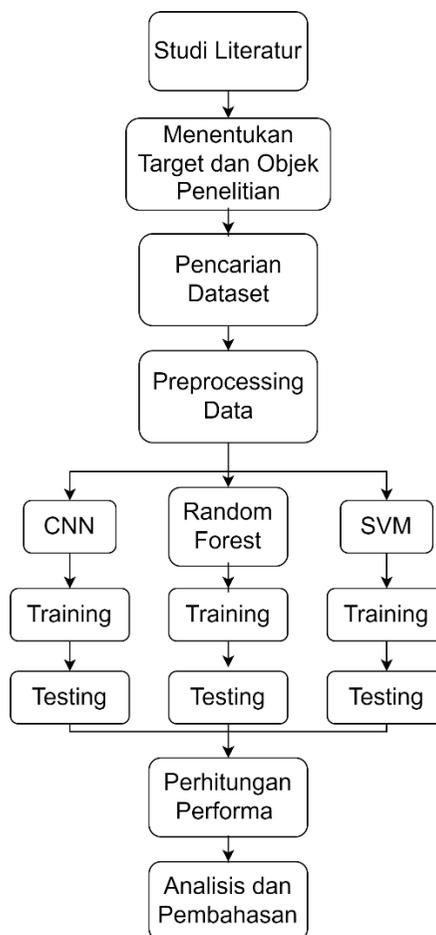
Pada tahap pra-penelitian, peneliti menentukan tema penelitian, masalah yang akan diteliti, mengumpulkan sumber referensi atau literatur seperti jurnal dan buku yang mendukung dalam penelitian, dan menentukan metode yang digunakan beserta batasan masalahnya. Kemudian peneliti mencari data yang sesuai dengan tema penelitian sebagai objek penelitian.

Pada tahap eksplorasi data, peneliti mencoba menguraikan karakteristik-karakteristik data yang digunakan untuk menentukan fungsi-fungsi tertentu yang akan digunakan. Setiap *dataset* akan memberikan parameter-parameter optimal yang berbeda terhadap algoritma klasifikasi *machine learning* (Santosa & Umam, 2018).

Pada tahap *data Training* dan *Validation*, peneliti akan mencari parameter-parameter terbaik untuk model yang akan digunakan berdasarkan hasil eksplorasi data dan *trial and error* untuk mendapatkan hasil terbaik. Kemudian data akan memberi hasil prediksi yang akan dianalisis pada tahap selanjutnya.

Pada tahap analisis hasil, peneliti akan merangkum hasil yang diperoleh dari metode-metode yang digunakan ke dalam bentuk tabel dan diagram, kemudian menyimpulkan hasilnya sebagai output dari penelitian ini.

2.4 Alur Penelitian



Gambar 7 Diagram alur penelitian.

Berdasarkan alur penelitian pada gambar 7, secara besar terbagi menjadi 4 tahapan yaitu Pra-penelitian, eksplorasi data, *set data Training, Validation*, dan analisis hasil. Pra-penelitian mencakup bagian studi literatur dan menentukan target dan objek penelitian, eksplorasi data mencakup bagian pengumpulan *dataset* dan menemukan metode yang akan digunakan. *Set data Training* dan *Validation* mencakup bagian melakukan *Training dan Validation* pada setiap metode yang digunakan. Dan yang terakhir analisis hasil merupakan bagian untuk melakukan analisis terhadap metode yang sudah dilakukan *Training* dan *Validation* kemudian menganalisis hasil dan melakukan pembahasannya.

2.5 Instrumen Penelitian

Instrumen yang digunakan dalam penelitian adalah *notebook* berspesifikasi *processor Intel® Core™ i5 7200U*, RAM 12GB, dan menggunakan sistem operasi *Windows 10 (localhost)*. Kemudian perangkat lunak yang digunakan untuk melakukan penelitian ini adalah:

- 1) *Google Collab pro.*
- 2) *Package machine learning* dari *Scikit.*