

**PREDIKSI JUMLAH PRODUKSI BUAH KELAPA SAWIT
MENGUNAKAN METODE *K-NEAREST NEIGHBOR*
(Studi Kasus : PT. Hardaya Inti Plantations di Kab. Buol)**

Oleh

SINTA LESTARI ASPA

T3115202

SKRIPSI

**Untuk memenuhi salah satu syarat ujian
Guna memperoleh gelar Sarjana**



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

2022

PENGESAHAN SKRIPSI

**PREDIKSI JUMLAH PRODUKSI BUAH KELAPA SAWIT
MENGUNAKAN METODE *K-NEAREST NEIGHBOR*
(Studi Kasus : PT. Hardaya Inti Plantations di Kab. Buol)**

Oleh

SINTA LESTARI ASPA

T3115202

SKRIPSI

Untuk memenuhi salah satu syarat ujian

Gunamemperoleh gelarSarjana

Dan telah disetujui oleh pembimbing pada bulan

Gorontalo, Juni2022

Pembimbing Utama



Haditsah Annur,M.kom
NIDN. 0919048404

Pembimbing Pendamping



Rofiq Harun,M.Kom
NIDN. 0919048404

PERSETUJUAN SKRIPSI

PREDIKSI JUMLAH PRODUKSI BUAH KELAPA SAWIT MENGUNAKAN METODE *K-NEAREST NEIGHBOR* (Studi Kasus : PT. Hardaya Inti Plantations di Kab. Buol)

Oleh

SINTA LESTARI ASPA

T3115202

Diperiksa Oleh Panitia Ujian Strata Satu (S1)

Universitas Ichsan Gorontalo

Gorontalo, 11 Juni 2022

1. Ketua Penguji
Irvan Abraham Salihi, M.Kom



2. Penguji
Sunarto Taliki, M.Kom



3. Penguji
Abd. Rahmat Karim Haba, M.Kom



4. Penguji
Haditsah Annur, M.Kom



5. Penguji
Rofiq Harun, M.Kom



Mengetahui :

Dekan Fakultas Ilmu Komputer


Jorry Karim, S.Kom. M.Kom
NIDN 0918077302

Ketua Program Studi


Sudirman Panna, S.Kom, M.Kom
NIDN 0922099101

PERNYATAAN SKRIPSI

Dengan saya menyatakan bahwa

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

Gorontalo, 11 Juni 2022



membuat pernyataan,

Sinta Lestari Aspa

ABSTRACT

SINTA LESTARI ASPA. T3115202. THE PREDICTION OF THE AMOUNT OF OIL PALM PRODUCTION USING THE K-NN METHOD

PT. Hardaya Inti Plantations is a company engaged in private oil palm plantations located in the Buol district. This company has been operating since 1995 with a total area up to 13,575 Ha. The production area is 13,072 Ha. This study discusses the prediction of the amount of oil palm production using the K-Nearest Neighbor method. The amount of oil palm production is uncertain every month due to several factors, namely rainfall, land area, and the number of trees. Therefore, an analysis is required to determine the amount of oil palm production in the following month using the K-Nearest Neighbor method. The object of this study is the amount of oil palm production. The data used is the historical data of the total production of oil palm plantations within 2016-2021 using the K-Nearest Neighbor method. The accuracy value with the best results is obtained using $K=3,5$ and 7 with a RMSE the lowest was obtained on the test $K=7$ with a value of 327,73.

Keywords: *K-Nearest Neighbor, oil palm production, RMSE*



ABSTRAK

SINTA LESTARI ASPA. T3115202. PREDIKSI JUMLAH PRODUKSI BUAH KELAPA SAWIT MENGGUNAKAN METODE K-NN

PT.Hardaya Inti Plantations Merupakan perusahaan yang bergerak dibidang perkebunana swasta kelapa sawit yang berada di kabupaten Buol. Perusahaan ini mulai beroperasi pada tahun 1995 sampai saat ini dengan total luas areal hingga 13.575 Ha dan luas areal yang menghasilkan 13.072 Ha. Penelitian ini membahas tentang prediksi jumlah produksi buah kelapa sawit menggunakan metode K-Nearest Neighbor, jumlah produksi buah kelapa sawit yang tidak menentu tiap bulannya karena disebabkan beberapa faktor yaitu curah hujan , luas lahan dan jumlah pokok . Oleh karena itu dibutuhkan suatu analisis untuk mengetahui jumlah produksi buah kelapa sawit pada bulan berikutnya dengan menggunakan metode K-Nearest Neighbor. Objek penelitian ini adalah jumlah produksi buah kelapa sawit data yang digunakan adalah data histori dari jumlah produksi tanaman kelapa sawit tahun 2016-2021 dengan menggunakan metode K-Nearest Neighbor. Didapatkan nilai akurasi dengan hasil terbaik menggunakan K=3,5 dan 7 dengan nilai RMSE nilai terendah didapatkan pada pengujian K=7 dengan nilai 327,73



Kata Kunci : *K-Nearest Neighbor*, produksi buah kelapa sawit, RMSE

KATA PENGANTAR

Bismillahirrahmanirrahim

Puji syukur penulis panjatkan kehadiran ALLAH SWT, karena atas berkat dan rahmat-Nya hingga dapat menyelesaikan Skripsi ini dengan judul **“Prediksi jumlah produksi buah kelapa sawit menggunakan metode *K-Nearest Neighbor*”** (studi kasus : PT. Hardaya Inti Plantations di Kab.Buol) untuk memenuhi salah satu syarat untuk memperoleh gelar Sarjana (S1) Program Studi Teknik Informatika Universitas Ichsan Gorontalo.

Penulis menyadari bahwa tanpa bantuan dan bimbingan dari berbagai pihak, Skripsi ini tidak dapat penulis selesaikan. Oleh karena itu penulis menyampaikan terimakasih kepada:

1. Ibu Dr. Hj. Juriko Abdussamad, M.Si, selaku Ketua Yayasan Pengembangan Ilmu Pengetahuan dan Teknologi (YPIPT) Ichsan Gorontalo;
2. Bapak Dr. Abdul Gaffar La Tjokke, M.Si, selaku Rektor Universitas Ichsan Gorontalo;
3. Bapak Jorry Karim, S.Kom, M.Kom, selaku Dekan Fakultas Ilmu Komputer Universitas Ichsan Gorontalo;
4. Bapak Sudirman Melangi, S.Kom, M.Kom, selaku Wakil Dekan I Bidang Akademik Fakultas Ilmu Komputer Universitas Ichsan Gorontalo;
5. Ibu Irma Surya Kumala, S.Kom, M.Kom, selaku Wakil Dekan II Bidang Adminidstrasi Umum dan Keuangan Fakultas Ilmu Komputer Universitas Ichsan Gorontalo;
6. Bapak Sudirman S. Panna, S.Kom, M.Kom, selaku Ketua Program Studi Teknik Informatika;
7. Ibu Haditsah Annur, M.Kom, selaku Pembimbing I yang telah membimbing penulis selama mengerjakan Skripsi ini;
8. Bapak Rofiq Harun, M.Kom, selaku Pembimbing II yang telah membimbing penulis selama mengerjakan Skripsi ini;

9. Bapak dan Ibu Dosen Universitas Ichsan Gorontalo yang telah mendidik dan membimbing penulis dalam mengerjakan Skripsi ini;
10. Teristimewa Kepada keluarga besar penulis, Ayah dan Ibu yang sudah memberikan kasih sayang kepada penulis, selalu memberikan doa, dan dukungan yang terbaik untuk penulis dari kecil hingga sudah sebesar ini.

Semoga beliau-beliau diatas mendapatkan imbalan yang lebih besar dari Allah SWT, melebihi apa yang beliau-beliau berikan kepada penulis. Saran dan kritik, penulis harapkan dari dewan penguji dan semua pihak untuk penyempurnaan penulisan skripsi lebih lanjut. Semoga Skripsi ini dapat bermanfaat bagi pihak yang berkepentingan.

Gorontalo, 11 Juni 2022

Penulis

DAFTAR ISI

HALAMAN SAMPUL	
HALAMAN PENGESAHAN.....	i
HALAMAN PERSETUJUAN	ii
HALAMAN PERNYATAAN.....	iii
ABSTRACT	iv
ABSTRAK	v
KATA PENGANTAR.....	vi
DAFTAR ISI.....	viii
DAFTAR GAMBAR.....	xi
DAFTAR TABEL	xii
BAB I PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Identifikasi Masalah	3
1.3 Rumusan Masalah	4
1.4 Tujuan Penelitian.....	4
1.5 Manfaat Penelitian.....	4
1.5.1 Manfaat Teoritis.....	4
1.5.2 Manfaat Praktis	4
BAB II LANDASAN TEORI	5
2.1 Tinjauan Studi	5
2.2 Tinjauan Pustaka	7
2.2.1 Kelapa Sawit.....	7
2.2.2 Produksi Kelapa Sawit.....	8
2.2.3 Prediksi	12
2.2.4 Data Mining.....	14
2.2.5 <i>K-Nearest Neighbor(K-NN)</i>	17
2.2.6 Contoh Penerapan Metode <i>K-Nearest Neighbor</i>	18
2.2.7 <i>Root Mean Square Error (RMSE)</i>	20
2.3 Kerangka Pikir	21
BAB III METODE PENELITIAN	22

3.1 Jenis, Metode, Subjek, Objek, dan Lokasi Penelitian	22
3.2 Pengumpulan Data.....	22
3.3 Pemodelan	24
3.3.1 Perhitungan <i>K-Nearest Neighbor</i>	24
3.3.2 Hasil Prediksi.....	25
3.3.3 Akurasi.....	25
BAB IV HASIL PENELITIAN.....	26
4.1 Model Usulan Penelitian.....	26
4.2 Dataset Penelitian	27
4.3 Pembagian Data.....	29
4.4 Pembentukan Model K-NN.....	32
4.5 Perhitungan K-NN.....	32
4.6 Hasil Prediksi.....	40
BAB V HASIL DAN PEMBAHASAN	42
5.1 Hasil Pemodel Dengan Rapid Miner.....	42
5.1.1 Tahapan Pemodelan Rapid Miner	42
5.1.1.1 Pengaturan Atribut Variabel.....	43
5.1.1.2 Pembagian Data Set.....	44
5.1.1.3 Properti K-NN	44
5.1.1.4 Performase.....	45
5.1.1.5 Hasil Prediksi.....	45
5.2 Pengujian Metode K-NN.....	46
5.2.1 RMSE (<i>Root Mean Square Error</i>).....	46
BAB VI PENUTUP	50
6.1 Kesimpulan.....	50
6.2 Saran	50
DAFTAR PUSTAKA.....	51
LAMPIRAN	
Lampiran 1. Hasil Prediksi Menggunakan Rapid Minner	
Lampiran 2. Surat Keterangan Penelitian	
Lampiran 3. Surat Rekomendasi Bebas Pustaka	

Lampiran 4. Hasil Uji Turnitin

Lampiran 5. Daftar Riwayat Hidup

DAFTAR GAMBAR

Gambar 2.1	Pertemuan disiplin ilmu dalam data mining.....	14
Gambar 2.2	Arsitektur sistem data mining.....	16
Gambar 2.3	Kerangka Pikir.....	21
Gambar 3.1	Model <i>K-Nearest Neighbor</i>	24
Gambar 4.1	Bagan Alir Model Usulan.....	26
Gambar 5.1	Memulai Rapid Miner	42
Gambar 5.2	Pemodelan K-NN Pada Rapid Miner	42
Gambar 5.3	Pengaturan atribut Variabel.....	43
Gambar 5.4	Split Validation.....	44
Gambar 5.5	Properti K-NN	44
Gambar 5.6	Performanse	45
Gambar 5.7	Hasil Prediksi.....	45
Gambar 5.8	Hasil Pengujian Dengan RMSE	48

DAFTAR TABEL

Tabel 1.1 Data Rekap Produksi Buah Kelapa Sawit.....	2
Tabel 2.1 Penelitian Terkait	5
Tabel 2.2 Contoh Data <i>Training</i>	18
Tabel 2.3 Perhitungan Kuadrat Jarak Data Training Dengan Data Testing.....	19
Tabel 2.4 Penentuan 3 Tetangga Terdekat Dari Data Testing	19
Tabel 3.1 Atribut Data	23
Tabel 4.1 Hasil Pengumpulan Data	27
Tabel 4.2 Data Training	29
Tabel 4.3 Data Testing.....	32
Tabel 4.4 Data Training	33
Tabel 4.5 Data Testing.....	33
Tabel 4.6 Hasil Perhitungan Jarak	34
Tabel 4.7 Peringkat Jarak.....	36
Tabel 4.8 Nilai $K=3$	38
Tabel 4.9 Nilai $K=5$	38
Tabel 4.10 Nilai $K=7$	39
Tabel 4.11 Perbandingan Data Aktual dengan Data Prediksi Untuk Masing-Masing Nilai K	40
Tabel 4.12 Hasil Prediksi dengan $K=3$	40
Tabel 4.13 Hasil prediksi dengan $K=5$	41
Tabel 4.14 Hasil Prediksi dengan $K=7$	41
Tabel 5.1 RMSE untuk Prediksi $K=3$	47
Tabel 5.2 RMSE untuk Prediksi $K=5$	47
Tabel 5.3 RMSE untuk Prediksi $K=7$	48
Tabel 5.4 Interpretasi Nilai RMSE	49

BAB I

PENDAHULUAN

1.1. Latar Belakang

Indonesia merupakan Negara dengan sektor tanaman perkebunan yang banyak di dominasi oleh tanaman perkebunan yakni kelapa sawit, kelapa sawit merupakan komoditas tanaman perkebunan yang memiliki peran strategis dalam menunjang pembangunan ekonomi di Indonesia[1]. Program ini sejalan dengan kebijakan yang di ambil oleh Pemerintah Indonesia dalam melaksanakan pembangunan berkelanjutan dengan memprioritaskan keseimbangan antara aspek sosial, ekonomi dan lingkungan, yang diatur dalam Rencana Pembangunan Jangka Menengah Nasional (RPJMN, 2020-2024). Seperti halnya pada wilayah yang ada di Provinsi Sulawesi Tengah yakni Kabupaten Buol yang merupakan salah satu daerah penghasil minyak kelapa sawit terbesar di Provinsi Sulawesi Tengah saat ini, dengan hasil produksi 209.528,76 Ton pada tahun 2013[2].

PT. Hardaya Inti Plantations merupakan perusahaan perkebunan swasta perkebunan kelapa sawit yang ada di Kabupaten Buol. Perusahaan ini Mulai beroperasi pada tahun 1995 sampai sekarang dengan total luas areal hingga 13.575 ha dan luas areal yang menghasilkan 13.072 ha (DataPT.HIP, 2020).Jumlah produksi kelapa sawit masih tidak menentu tiap bulannya karena sering mengalami peningkatan dan penurunan pada jumlah produksinya, dimana banyak sedikitnya hasil produksi tanaman buah kelapa sawit dipengaruhi oleh tingkat penggunaanfaktor produksi. Salah satu faktor yang turut menentukan tingkat hasil produksi buah kelapa sawit adalah Curah hujan yang kurang dapat dapat berdampak pada perkembangan bunga pada tanaman buah kelapa sawit, serta meningkatny keguguran, tanda gagal,atau buah busuk, serta produktivitasnya rendah dan perbungaan panjang hingga 9 bulan lamanya. Luas lahan,keberadaan luas lahan sangat penting dalam menunjang kegiatan produksi hasil pertanian. dan jumlah pokok yang kurang dapat mempengaruhi produktivitas kelapa sawit menjadi menurun. Dalam menargetkan jumlah hasil produksi buah kelapa sawit berikutnya perusahaan harus melakukan prediksi terlebih dahulu agar jumlah produksi yang diinginkan dapat terpenuhi.

Berikut adalah sampel data yang di peroleh dari pihak responden perusahaan kelapa sawit PT. Hardaya Inti Plantations dari tahun 2016- 2021.

Tabel 1.1 Data RekapHasil Produksi Buah Kelapa Pada Divisi III Tahun 2016 – 2021.

BULAN	PRODUKSI (Ton)TAHUN 2016-2021					
	2016	2017	2018	2019	2020	2021
Januari	3,551,130	5,025,150	2,324,300	4,401,800	4,874,180	4,417,820
Februari	3,098,590	3,861,600	1,827,540	3,133,540	4,207,790	3,745,230
Maret	2,905,590	4,034,010	2,385,655	3,494,660	3,931,300	3,946,400
April	2,618,320	4,529,840	3,281,659	2,745,090	4,326,110	3,846,040
Mei	2,268,110	4,681,830	4,366,220	2,456,013	3,369,090	2,278,360
Juni	1,825,460	3,629,540	4,308,350	2,208,365	3,927,060	4,020,790
Juli	1,668,800	5,316,800	7,031,950	3,562,354	2,264,960	3,075,660
Agustus	2,778,820	4,541,430	6,560,950	3,968,660	2,981,940	2,772,850
September	6,439,000	4,425,880	5,533,410	4,597,630	3,579,930	3,224,432
Oktober	6,474,030	4,156,600	6,179,635	5,912,060	4,237,770	3,535,615
November	6,951,690	3,991,600	5,393,710	6,075,710	4,772,620	3,812,423
Desember	5,893,380	2,886,216	4,671,290	4,918,360	4,656,060	3,639,928
TOTAL	46,472,720	51,080,496	53,864,669	47,474,252	46,128,810	42,315,548

(Sumber: PT. HARDAYA INTI PLANTATIONS Tahun 2021)

Data diatas merupakan data sekunder yang di peroleh dari pihak responden bagian Mill Manager PT. HIP yang memuat tentang data Rekap Produksi dari tahun 2016 sampai 2021 pada DIV III. Dapat dilihat dari tabel data diatas bisa disimpulkan bahwa produksi buah kelapa sawit pada tahun 2017-2018 jumlah produksi sangat banyak, sedangkan pada tahun 2016, 2019, 2020 dan 2021 jumlah produksi buah kelapa sawit sedikit. Yang berarti jumlah produksi buah kelapa sawit dapat diterangkan bersama sama oleh variable luas lahan, jumlah pokok dan curah hujan. Seperti yang kita ketahui bahwa faktor-faktor produksi tersebut sangat berpengaruh nyata terhadap jumlah produksi buah kelapa sawit.

Menyikapi masalah diatas maka dibutuhkan suatu metode komputasi yang tepat untuk mempermudah dalam pengambilan keputusan dalam memprediksi jumlah produksi buah kelapa sawit agar kedepannya mendapatkan hasil produksi yang di inginkan.

Salah satu metode komputasi yang sering digunakan dalam kasus menentukan prediksi jumlah produksi adalah metode *K-Nearest Neighbor* . Pemilihan metode

K-Nearest Neighbor adalah suatu metode yang menggunakan algoritma *supervised* dimana hasil dari query instance yang baru diklasifikasikan berdasarkan mayoritas dari label class pada *K-Nearest Neighbor*. Tujuan dari algoritma *K-Nearest Neighbor* adalah mengklasifikasikan objek baru berdasarkan atribut dan training data. Algoritma *K-Nearest Neighbor* bekerja berdasarkan jarak terpendek dari query instance ke training data untuk menentukan *K-Nearest Neighbor*-nya.[3].

Berikut adalah penelitian yang membahas mengenai produktifitas kelapa sawit dilakukan oleh EmaAgasta, dkk[4], mengenai Prediksi Jumlah Produksi Kelapa Sawit dengan Menggunakan Metode *Extreme Learning Machine* (ELM) yang menggunakan 4 kriteria seperti umur tanam, jumlah pokok/jumlah pohon, luas lahan dan hasil produksi. Berdasarkan hasil pengujian maka model prediksi yang digunakan memiliki kinerja yang cukup bagus, karena nilai RMSE berada diantara rentang 20 - 50%..

Berdasarkan hasil pengamatan dari sumber referensi diatas, terdapat beberapa kemiripan pada kriteria yang digunakan. namun memiliki perbedaan terhadap metode yang digunakan dengan penelitian ini. metode yang akan di gunakan dalam penelitian ini yaitu metode *K-Nearest Neighbor* Selanjutnya untuk kriteria yang akan digunakan dalam penelitian ini yakni kriteria , Kriteria tersebut selain diperoleh dari hasil wawancara dengan pihak responden PT. Hardaya Inti Plantations, kriteria diatas juga merupakan hasil pengamatan maupun perbandingan penulis terhadap sumber-sumber refensi yang pernah dilakukan oleh peniliti-peneliti sebelumnya.

Dari uraian diatas, maka peneliti ingin melakukan penelitian dengan judul: **“Prediksi Jumlah Produksi Buah Kelapa Sawit Menggunakan Metode *K-Nearest Neighbor*”** Studi Kasus PT.Hardaya Inti Plantations di Kabupaten Buol.

1.2 Identifikasi Masalah

Berdasarkan latar belakang di atas, maka identifikasi masalahnya adalah : Perusahaan kelapa sawit PT. Hardaya Inti Plantations sering mengalami kendala dimana sering terjadi peningkatan dan penurunan hasil produksi buah kelapa sawit

pada setiap tahunnya maka perlu dilakukan proses prediksi dengan menerapkan metode *K-Nearest Neighbor*

1.3 Rumusan Masalah

Bagaimana kinerja metode *K-Nearest Neighbor* dalam melakukan prediksi jumlah produksi buah kelapa sawit?

1.4 Tujuan Penelitian

Penelitian/study ini dilakukan dengan tujuan demi mengetahui kinerja metode *K-Nearest Neighbor* dalam melakukan prediksi jumlah produksi buah kelapa sawit pada perusahaan kelapa sawit milik PT. Hardaya Inti Plantations di Buol.

1.5 Manfaat Penelitian

1.5.1 Manfaat teoritis

Penelitian ini diharapkan dapat memberikan sumbangsi dan masukan terhadap pengembangan ilmu pengetahuan dibidang teknologi computer pada umumnya dan tentang penerapan metode *K-Nearest Neighbor* dalam melakukan prediksi.

1.5.2 Manfaat praktis.

Sebagai salah satu bahan kajian bagi semua elemen-elemen ataupun unsur unsur yang terlibat dalam perancangan metode *K-Nearest Neighbor* dalam melakukan prediksi jumlah produksi buah kelapa sawit di Pt. Hardaya Inti Plantations.

BAB II

LANDASAN TEORI

2.1 Tinjauan Studi

Berdasarkan Penelitian yang sebelumnya pernah dilakukan oleh beberapa peneliti yang menjadi dasar acuan penulis untuk mengembangkan teori, misalnya penelitian yang dilakukan oleh:

Tabel 2.1 Penelitian Terkait

No	PENELITI/ TAHUN	JUDUL	METODE	HASIL
1.	Irma Surya Kumala Idkk. 2019 [5]	Metode K-Nearest Neighbor untuk Memprediksi Harga Eceran Beras di Pasar Tradisional Gorontalo	<i>K-Nearest Neighbor</i>	Berdasarkan penelitian sistem prediksi harga eceran beras di pasar tradisional Gorontalo menghasilkan sistem prediksi harga eceran beras penjual dapat menyusun target penjualan dan pembeli dapat mengantisipasi ketidakpastian harga kebutuhan beras. Sistem prediksi harga eceran beras di pasar tradisional Gorontalo menggunakan metode K-Nearest Neighbor mendapatkan tingkat akurasi sebesar 80%..
2.	Natcha Kwintarini Suparman 2021[6].	Prediksi Barang Keluar TB.Wijaya Bangunan menggunakan Algoritma <i>K-Nearest Neighbor Regresion</i> dengan RStudio	<i>K-Nearest Neighbor</i>	Berdasarkan hasil penelitian yang telah dilaksanakan maka didapat kesimpulan hasil prediksi penjualan bahan bangunan yang paling sering keluar selama 9 periode perminggu di tahun 2020 sebanyak 5 produk bahan bangunan dari 60 produk bahan bangunan, yaitu cat yoko, mangkoklistrik, kunci kaca faster, lis 20cm, dan klem selang. Serta ditampilkan hasil

				<p>visualisasi berupa grafik plot hasil prediksi yang telah didapat dengan menggunakan Shiny <i>framework</i> untuk mempermudah dalam menampilkan hasil prediksi. Hasil evaluasi perhitungan tingkat akurasi dengan menggunakan metode RMSE (<i>Root Means Square Error</i>) dengan hasil bahwa Engsel Sendok Cobra Sherlock mendapat hasil evaluasi paling kecil dengan nilai <i>error</i> 3.55 yang berarti memiliki hasil akurasi terbaik. <i>Package</i> “<i>tsfkn</i>” pada RStudio juga cukup sederhana digunakan untuk melakukan prediksi dengan algoritma <i>time series k-nearest neighbor regression</i> karena <i>package</i> ini cukup menggunakan satu <i>function</i> atau kode-kode yang disusun untuk melakukan suatu tugas dengan menggabungkan beberapa perintah dalam satu kode pemrograman pada RStudio.</p>
3.	Esty Purwaningsih 2021[7].	Penerapan <i>K-Nearest Neighbor</i> untuk Klasifikasi Tingkat Kelulusan Pada Siswa	<i>K-Nearest Neighbor</i>	<p>Metode <i>K-Nearest Neighbor</i> (KNN) dapat memprediksi tingkat kelulusan siswa. Metode <i>k-nearest neighbor</i> (KNN) yang diproses dengan <i>tools rapidminer</i> 9.0 didapatkan rata-rata akurasi sebesar 96,49%. Tingkat akurasi pengujian model kelulusan siswa dengan menggunakan algoritma <i>K-Nearest Neighbor</i> (K-NN) juga dipengaruhi oleh jumlah klastering data. Sistem yang digunakan dengan <i>rapid</i></p>

				<i>miner</i> dapat membantu dalam menentukan <i>kriteria</i> data siswa lulus mencapai KKM dan tidak lulus mencapai KKM sehingga membantu dalam pengambilan keputusan.
--	--	--	--	--

2.2 Tinjauan Pustaka

2.2.1 Kelapa Sawit

Kelapa sawit (*elaeis guineensis jacq*) merupakan salah satu tanaman yang berasal dari Benua Afrika. Kelapa sawit terdapat negara beriklim hutan hujan tropis seperti negara kamerunm Pantai Gading, Ghana, Liberia, Nigeria, Sierra Leono, Angola dan Kongo. Minyak kelapa sawit tersebut digunakan penduduk setempat untuk memasak dan bahan untuk kecantikan, selain dari pada itu, buah kelapa sawit juga dapat diolah menjadi minyak nabati. Minyak kelapa sawit mengandung *karotenbid* yang lumayan tinggi, *karotenbid* merupakan pigmen yang menghasilkan warna merah, dan terdapat asam jenuh palmitat sebagai komponen utamanya yang mengakibatkan minyak bertekstur kental-semi padat dan jika tumbuh di daerah beriklim sedang maka akan menjadi lemak padat [8]

Tanaman kelapa sawit juga merupakan tanaman hutan yang berasal dari kawasan Amerika Selatan yakni Brazil dimana spesies tanaman ini umumnya banyak banyak ditemukan di hutan Brazil dibanding dengan di Afrika, Pada ekspetasinya tanaman ini (kelapa sawit) justru tumbuh subur diluar daerah asalnya, seperti di Malaysia, Thailand, Papua Nugini dan di Indonesia bahkan mampu memberikan hasil produksi per hektar yang lebih tinggi ketimbang produksi pada negara asal tumbuhan ini. Bagi Indonesia, kelapa sawit memiliki peran penting dalam pembangunan perkebunan nasional, dan mampu menciptakan lapangan kerja yang mengacu pada aspek kesejahteraan masyarakat. Sehingga saat ini Indonesia merupakan salah satu negara produsen penghasil minyak sawit di dunia internasional dan juga miliki luas areal perkebunan terluas di dunia [9].

Berikut ini merupakan manfaat dari tanaman kelapa sawit yang mana industri ini menghasilkan manfaat utama sebagai minyak masak (minyak goreng), minyak industri, dan bahan bakar (biodiesel). Selain itu, kelapa sawit juga merupakan bahan baku untuk industri sabun, lilin, pembuatan lembaran-lembaran timah, dan industri dalam bidang kosmetik kecantikan. Usaha perkebunan kelapa sawit merupakan sebuah potensi bisnis yang sangat menguntungkan melihat pemanfaatannya yang baik dari segi industri makanan hingga industri kimia, yang diantaranya industri mentega, *shortening*, cokelat, bahan aditif, es krim, pakan ternak serta masih banyak lagi. Selain itu dalam industri kulit juga membutuhkan bahan baku dari kelapa sawit yang berguna untuk membuat bahan kulit menjadi halus dan lentur dan tahan terhadap tekanan tinggi ataupun temperatur tinggi. Pada industri perak, hasil olahan kelapa sawit berguna pada proses *cold rolling and fluning agent*. Dan pada industri logam juga menggunakan bahan kelapa sawit yakni sebagai bahan untuk pemisah material dan tembaga [8]

2.2.2 Produksi Kelapa Sawit

Perkembangan industri pada produksi minyak kelapa sawit di Negara beriklim tropis telah didorong oleh potensi produktivitas yang memang sangat tinggi. Soalnya, kelapa sawit memberikan hasil tinggi minyak persatuan luas dibandingkan dengan tanaman lainnya. Selain itu hasil produksi tanaman ini ternyata menghasilkan dua jenis minyak, yaitu minyak kelapa sawit dan minyak sawit kernel (inti). Kedua jenis minyak yang dihasilkan dari kelapa sawit tersebut sangat diminati oleh pasar global[8]. dari hasil pencapaian produksi umumnya kelapa sawit di Indonesia pada awal tahun 2004-2008 tercatat sebesar 75,54 juta ton tandan buah segar (TBS) atau 40,26% dari total produksi kelapa sawit dunia[9].Produksi minyak sawit dan inti sawit pada tahun 2018 tercatat sebesar 48,68 juta ton, yang terdiri dari 40,57juta ton *crude palm oil* (CPO) dan 8,11 juta ton *palm kernel oil* (PKO). Jumlah produksi tersebut berasal dariPerkebunan Rakyat sebesar 16,8 juta ton (35%), Perkebunan Besar Negara sebesar 2,49 juta ton (5%), danPerkebunan Besar Swasta sebesar 29,39 juta ton (60%)[1].Sedangkan hasil produksi minyak

kelapa sawit di Provinsi Sulawesi Tengah pada tahun 2013, dengan hasil produksi 209.528,76 Ton[2].

Dari hasil produksi kelapa sawit yang dilihat pada tahun-tahun sebelumnya diatas, artinya tingkat keberhasilan hasil produksi tersebut di pengaruhi oleh banyak faktor, baik itu faktor dari luar maupun dari tanaman kelapa sawit itu sendiri. Pada umumnya faktor-faktor tersebut dibedakan menjadi faktor lingkungan, faktor genetis, dan faktor teknis-argonomis. Dalam menunjang pertumbuhan dan proses produksi kelapa sawit, faktor tersebut harus saling berkaitan dan saling mempengaruhi satu sama lainnya untuk mendapatkan capaian hasil produksi yang maksimal, diharapkan faktor tersebut selalu dalam kondisi optimal [9]. Faktor lingkungan tersebut merupakan faktor yang umum yang sangat mendominasi dalam hasil produksi kelapa sawit, uraian dari faktor tersebut sebagai berikut :

1) Faktor Iklim

Faktor ini sangat berpengaruh terhadap pertumbuhan dan produksi tandan kelapa sawit . yang dapat tumbuh dengan baik jika berada di daerah tropika basah di antara 12° LU – 12° LS pada ketinggian 0 – 500 m dpl. Pada daerah sekitar lintasan garis khatulistiwa kelapa sawit masih bisa menghasilkan buah pada ketinggian 1.300 m dpl. Berikut ini adalah unsur dari faktor iklim yang mempengaruhi :

a. Curah Hujan

Curah hujan optimum rata-rata yang diperlukan tanaman kelapa sawit adalah 2.000-2.500 mm/tahun dengan distribusi merata sepanjang tahun tanpa bulan kering(defisit air)yang berkepanjangan.Curah hujan yang merata dapat menurunkan penguapan dari tanah dan tanaman kelapa sawit.Namun, yang terpenting adalah tidak terjadi defisit air diatas 250 mm. Bila tanah dalam keadaan kering,akar tanaman sulit menyerap mineral dari dalam tanah. Oleh sebab itu,musim kemarau yang berkepanjangan cenderung akan menurunkan produksi. Daerah di Indonesia yang sering mengalami kekeringan adalah Lampung dan Jawa Barat, sedangkan Kalimantan Timur dan beberapalokasi lainnya hamper setiap 5-6 tahun sekali.

Umumnya daerah dengan curah hujan yang tinggi kadang menjadi masalah terutama jalan untuk transport, pembakaran, pemeliharaan, pemupukan, dan pencegahan erosi. Daerah di Indonesia seperti ini kebanyakan berada pada ketinggian diatas 500 m dpl, kecuali di beberapa lokasi pantai barat Sumatera. Sebagian besar perkebunan kelapa sawit komersil dibangun pada daerah yang mempunyai neraca air positif selama enam bulan atau lebih. Neraca air positif adalah kondisi curah hujan lebih besar dari pada evapotranspirasi di perkebunan.

b. Sinar Matahari

Tanaman kelapa memerlukan intensitas cahaya yang tinggi untuk berfotosintesis, kecuali saat kondisi tanaman masih juvenile di *pre-nursery*. Sinar matahari diperlukan untuk memproduksi karbohidrat dan memacu pembentukan bunga dan buah. Untuk itu, intensitas, kualitas, dan lama penyinaran amat berpengaruh. Lama penyinaran optimum yang diperlukan kelapa sawit antara 5-12 jam/hari. Beberapa daerah seperti Riau, Jambi dan Sumatera Selatan sering terjadi penyinaran matahari kurang dari 5 jam pada bulan-bulan tertentu. Penyinaran yang kurang menyebabkan berkurangnya asimilasi dan gangguan penyakit.

c. Suhu

suhu optimum yang dibutuhkan agar tanaman kelapa sawit dapat tumbuh dengan baik adalah 24-28°C. sementara itu, untuk produksi Tandan Buah Segar (TBS) yang tinggi, maka diperlukan suhu pada rata-rata tahunan berkisar 25-27°C. walaupun demikian, tanaman ini masih bisa tumbuh pada suhu terendah yakni 18°C dan suhu tertinggi yakni 32°C. pada suhu 15°C, pertumbuhan tanaman tersebut sudah mulai terhambat. Beberapa faktor yang mempengaruhi tinggi rendahnya suhu adalah lama penyinaran dan ketinggian lokasi. Makin lama penyinaran atau makin rendah suatu lokasi maka makin tinggi suhunya. Suhu berpengaruh terhadap masa pembungaan dan kematangan buah. Kelapa sawit yang ditanam pada ketinggian diatas 500 m dpl akan berbunga lebih lambat satu tahun dibandingkan dengan tanaman di dataran rendah.

d. Kelembapan Udara dan Angin

Faktor ini merupakan faktor penting dan penunjang pertumbuhan kelapa sawit. Kelembapan optimum bagi pertumbuhan kelapa sawit adalah 80%. Faktor-faktor yang mempengaruhi kelembapan adalah suhu, sinar matahari, lama penyinaran, curah hujan, dan *evapotranspirasi*. Kecepatan angin 5-6 km/jam sangatlah baik untuk membantu proses penyerbukan dari bunga kelapa sawit, angin yang kering menyebabkan penguapan lebih besar, mengurangi kelembapan dan dalam waktu lama mengakibatkan tanaman layu. Sementara itu, angin yang terlalu kencang dapat menjadikan tanaman yang baru tumbuh menjadi miring.

2) Faktor Tanah

Kelapa sawit dapat hidup ditanah mineral, gambut, dan pasang surut. Potensi pengembangan kelapa sawit di lahan gambut (organic) relatif sangat baik, sifat fisik tanah gambut diantaranya selalu tergenang air, dekomposisi bahan organik lambat, konsistensinya lemas, kepadatan massa rendah dan bersifat seperti spons. Drainase di lahan gambut biasanya diikuti oleh penyusutan massa dan penurunan muka tanah. Berikut ini sifat kimia tanah gambut[7] :

- a) Kandungan organik mentah sangat tinggi
- b) Asam humik dan fulfik tinggi
- c) Keasaman tanah (pH) 3-3,5
- d) Kandungan N tinggi
- e) Kandungan C/N dan KPK tinggi
- f) Kandungan unsur P, K, Mg, Cu, Zn, dan B dalam kondisi defisien.

Selain tanah gambut, jenis tanah yang potensial untuk kelapa sawit ialah tanah sulfat asam (pasang surut) dengan luas Kawasan di Indonesia mencapai kisaran 2 juta hektar, kelebihan dari tanah ini adalah lokasi berada di sekitar daerah pantai (pasang surut) dengan topografi datar. Sementara ini, kekurangan tanah pasang surut yaitu memiliki kandungan senyawa pirit tinggi (FeS) dan beresiko mengalami oksidasi. Kandungan NaCl dan MgCl sangat tinggi karena letaknya di daerah pasang surut air laut. Selain itu, air tanah sangat pekat dan pertanaman akan

mengalami *plasmolysis* karena itu, perlu ditemukan teknologi yang tepat untuk mengelola sistem drainase di tanah sulfat asam (Pasang Surut).

2.2.3 Prediksi

Prediksi adalah suatu proses memperkirakan secara sistematis tentang sesuatu yang paling mungkin terjadi di masa depan berdasarkan informasi masa lalu dan sekarang yang dimiliki, agar kesalahannya (selisih antara sesuatu yang terjadi dengan hasil perkiraan) dapat diperkecil. Prediksi tidak harus memberikan jawaban secara pasti kejadian yang akan terjadi, melainkan berusaha untuk mencari jawaban sedekat mungkin yang akan terjadi[10]. Pengertian Prediksi sama dengan ramalan atau perkiraan. Menurut kamus besar bahasa Indonesia, prediksi adalah hasil dari kegiatan memprediksi atau meramal atau memperkirakan nilai pada masa yang akan datang dengan menggunakan data masa lalu. Prediksi menunjukkan apa yang akan terjadi pada suatu keadaan tertentu dan merupakan input bagi proses perencanaan dan pengambilan keputusan[11].

Secara *Eksplisit*, pembahasan mengenai teori peramalan kebijakan sangatlah sedikit. Namun, secara implisit, peramalan kebijakan terkait menjadi satu dengan proses analisa kebijakan. Karena didalam menganalisa kebijakan, untuk menformulasikan sebuah rekomendasi kebijakan baru, maka diperlukan adanya peramalan-peramalan atau prediksi mengenai kebijakan yang akan diberlakukan dimasa yang akan datang. Namun, satu dari sekian banyak prosedur yang ditawarkan oleh para pakar Dunn, masih memberikan pembahasan tersendiri mengenai peramalan kebijakan[11].

Peramalan (*forecasting*) adalah suatu prosedur untuk membuat informasi factual tentang situasi sosial masa depan atas dasar informasi yang telah ada tentang masalah kebijakan. Ramalan mempunyai tiga bentuk utama: proyeksi, prediksi, dan perkiraan berikut pengetiannya[12].

1. Suatu proyeksi adalah ramalan yang didasarkan pada ekstrapolasi atas kecenderungan masa lalu maupun masa kini ke masa depan. Proyeksi

membuat pertanyaan yang tegas berdasarkan argument yang diperoleh dari metode tertentu dan kasus yang paralel.

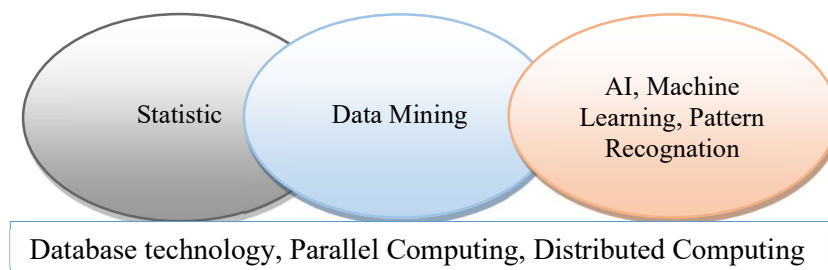
2. Sebuah prediksi adalah ramalan yang didasarkan pada asumsi teoritik yang tegas. Asumsi ini dapat berbentuk hukum teoretis (misalnya hukum berkurangnya nilai uang), proposisi teoritis (misalnya proposisi bahwa pecahnya masyarakat sipil diakibatkan oleh kesenjangan antara harapan dan kemampuan), atau analogi (misalnya analogi antara pertumbuhan organisasi pemerintah dengan pertumbuhan organisme biologis).
3. Suatu perkiraan (conjecture) adalah ramalan yang didasarkan pada penilaian yang informative atau penilaian pakar tentang situasi masyarakat masa depan.

Tujuan dari pada diadakannya peramalan kebijakan adalah untuk memperoleh informasi mengenai perubahan dimasa yang akan datang yang akan mempengaruhi terhadap implementasi kebijakan serta konsekuensinya[11]. Olehsebabnya, sebelum rekomendasi diformulasikan perlu adanya peramalan kebijakan sehingga akan diperoleh hasil rekomendasi yang benar-benar akurat untuk diberlakukan pada masa yang akan. Didalam memprediksi kebutuhan yang akan datang dengan berpijak pada masa lalu, dibutuhkan seseorang yang memiliki daya sensitifitas tinggi dan mampu membaca kemungkinan-kemungkinan dimasa yang akan datang. Permalan kebijakan juga diperlukan untuk mengontrol, dalam artian, berusaha merencanakan dan menetapkan kebijakan sehingga dapat memberikan alternatif-alternatif tindakan yang terbaik yang dapat dipilih diantara berbagai kemungkinan yang ditawarkan oleh masa depan. Masa depan juga terkadang banyak dipengaruhi oleh masa lalu. Dengan mengacu pada masa depan analisis kebijakan harus mampu menaksir nilai apa yang bisa atau harus membimbing tindakan di masa depan[13].

2.2.4 Data Mining

Data mining merupakan salah satu proses dimana pengumpulan data atau informasi penting yang tertuang dalam sebuah data besar, dimana data ataupun informasi tersebut dapat diidentifikasi ataupun diekstraksi menjadi sebuah

informasi yang bermanfaat untuk pengetahuan yang terhubung dengan berbagai database berskala besar[14]. Data mining atau disebut juga *Knowledge Discoveri in Database* (KDD) ialah proses pengambilan informasi yang tersembunyi, dimana informasi tersebut awalnya tidak diketahui, proses dalam KDD ini meliputi pendekatan secara teknis seperti pada *clustering*, klasifikasi yang merupakan irisan *artificial intelligence* (AI), *machine learning* (ML) dan Statistik [15].



Gambar 2.1 : Pertemuan disiplin ilmu dalam data mining [15]

Menurut Gatner Group[16] data mining adalah suatu proses menemukan hubungan yang berarti, pola dan kecenderungan dengan memeriksa dalam sekumpulan besar data yang tersimpan dalam penyimpanan, dengan menggunakan teknik pengenalan pola seperti teknik statistik dan matematika.

Beberapa karakteristik data mining adalah sebagai berikut [15] :

- 1) Data mining berkaitan erat dengan pola yang tersembunyi dan pola yang tidak dapat diketahui sebelumnya.
- 2) Data mining menggunakan data besar, yang dapat meningkatkan hasil tingkat kepercayaan jika data tersebut berskala besar.
- 3) Data mining juga berfungsi guna membuat keputusan kritis dalam hal strategi.

Seiring berkembangnya zaman data mining memiliki beberapa faktor yang mendorong perkembangan di antaranya[14] :

- 1) Adanya pertumbuhan kumpulan data dan informasi yang sangat cepat.

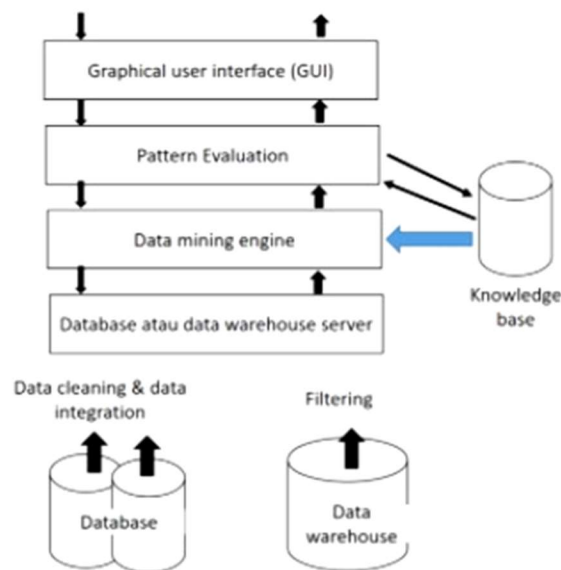
- 2) Hampir semua perusahaan datanya disimpan dalam *datawarehouse*, sehingga pihak perusahaan sudah memiliki akses ke dalam *databasenya*.
- 3) Mulai meningkatnya penggunaan akses data melalui navigasi web dan internet.
- 4) terdapat persaingan yang kompetitif dalam bidang bisnis yang bertujuan untuk meningkatkan penguasaan pasar di era globalisasi sekarang ini.
- 5) pesatnya penggunaan *data mining* dalam perkembangan teknologi perangkat lunak.

Kegunaan data mining secara garis besarnya terbagi menjadi dua yakni : *deskriptif* dan *prediktif*. Data mining dalam artian *deskriptif* merupakan pola-pola yang di temukan yang digunakan untuk menjelaskan karakteristik data, sedangkan pada artian *prediktif* yakni data mining yang digunakan untuk menemukan model-model pengetahuan dalam melakukan prediksi. Berdasarkan fungsional data mining tersebut pembagiannya meliputi klasifikasi, klastering, regresi, *association rule*, *anomaly detection*, dan *summarization* atau disebut penunjang informasi dan penguatan keputusan [15].

Data mining ialah proses dari penggalian sebuah pengetahuan yang ada pada data-data yang tertuang dalam jumlah maupun ukuran yang besar dan terimpan dalam basis data (*database*). Data *warehouse* pada arsitektur sistem data mining memiliki komponen-komponen utama yakni :

- 1) Adanya tempat penyimpanan informasi lainnya atau disebut data *warehouse*.
- 2) Data *warehouse server* merupakan sebuah komponen yang mempunyai tugas Dalam melakukan pengambilan data relevan berdasarkan data kebutuhan penggunaanya.
- 3) Basis pengetahuan, komponen ini merupakan domain *knowledge* yang berguna untuk melakukan proses pencarian atau mengevaluasi terhadap pola-pola.

- 4) Data mining *engine*, merupakan bagian terpenting dalam arsitektur data mining yang memuat modul-modul fungsional pada data mining seperti ; klasifikasi,asosiasi,*cluster* dan seterusnya.
- 5) Modul evaluasi pola, ialah modul data mining *engine* dalam tahapan penarikan pola-pola yang saling berinteraksi. Modul evaluasi pola ini menggunakan *threshold* agar memperoleh dan menyaring pola yang akan diperoleh.



Gambar 2.2 : Arsitektur sistem data mining [15]

- 6) *User Interface* atau antarmuka pengguna, modul ini merupakan komponen yang digunakan untuk berkomunikasi/berinteraksi dengan penggunanya. Informasi yang disediakan akan berfokus pada pencarian dan eksplorasi data mining.

2.2.5 *K-Nearest Neighbor*

K-Nearest Neighbour (K-NN) adalah suatu metode yang menggunakan algoritma *supervised* dimana hasil dari sampel uji yang baru diklasifikasikan berdasarkan mayoritas dari kategori pada K-NN. Tujuan dari algoritma ini adalah mengklasifikasi objek baru berdasarkan atribut dan sampel latih. Pengklasifikasian tidak menggunakan model apapun untuk dicocokkan dan hanya berdasarkan pada

memori. Diberikan titik uji, akan ditemukan sejumlah K objek (titik training) yang paling dekat dengan titik uji. Klasifikasi menggunakan *voting* paling banyak di antara klasifikasi dari K objek. Algoritma K-NN menggunakan klasifikasi ketetanggaan yang berfungsi sebagai nilai prediksi dari sample uji yang baru. Dekat atau jauhnya tetangga biasanya dihitung berdasarkan jarak *Euclidean*[16].

Algoritma metode KNN sangatlah sederhana, bekerja berdasarkan jarak terdekat dari sampel uji ke sampel latih untuk menentukan KNN-nya. Untuk merepresentasikan masing-masing fitur dari data, sampel latih diproyeksikan ke ruang berdimensi banyak. Berdasarkan klasifikasi sampel latih ruang ini dibagi menjadi bagian-bagian. Titik pada ruang ini ditandai dengan kelas c, jika kelas c merupakan klasifikasi yang paling banyak ditemui pada K sebuah tetangga terdekat dari titik tersebut

Tujuan dari algoritma k-NN adalah mengklasifikasikan objek baru berdasarkan atribut dan training data. Algoritma k-NN bekerja berdasarkan jarak terdekat dari query instance ke training data untuk menentukan k-NN-nya. Salah satu cara untuk menghitung jarak dekat atau jauhnya tetangga menggunakan metode euclidean distance. Euclidean Distance sering digunakan untuk menghitung jarak. Euclidean Distance berfungsi menguji ukuran yang bisa digunakan sebagai interpretasi kedekatan jarak antara dua objek, di bawah ini merupakan rumus Euclidean Distance:

Keterangan:

$$d(x_i, x_j) = \sqrt{\sum_{r=1}^n (a_r(x_i) - a_r(x_j))^2}$$

d (xi, xj): Jarak Euclidean (Euclidean Distance).

(xi): record ke-i

(xj): record ke- j

(ar) : data ke-r

i, j:1,2,3,...n

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

Langkah-langkah untuk menghitung *Metode Algoritma K-Nearest Neighbor*:

1. Menentukan parameter K (jumlah tetangga paling dekat).
2. Menghitung kuadrat jarak *Euclid (query instance)* masing-masing objek terhadap data sampel yang diberikan.
3. Kemudian mengurutkan objek-objek tersebut kedalam kelompok yang mempunyai jarak *Euclid* terkecil.
4. Mengumpulkan kategori Y (*Klasifikasi K-Neares Neighbor*).
5. Dengan menggunakan kategori *Nearest Neighbor* yang paling mayoritas, maka dapat diprediksikan nilai *query instance* yang telah dihitung

2.2.6 Contoh Penerapan Metode K-Nearest Neighbor

Diberikan data Training berupa dua atribut X1 dan X2 serta Y untuk mengklasiikasikan sebuah data apakah tergolong Baik atau Buruk , berikut ini adalah contoh datanya :

Tabel 2.2 Contoh Data *Training*

X1	X2	Y
7	7	Buruk
7	4	Buruk
3	4	Baik
1	4	Baik

Kita diberikan data baru yang akan kita klasifikasikan, yaitu **X1= 3** dan **X2= 7**,tentukan nilai Y! Langkah penyelesaian:

1. Tentukan parameter K = jumlah tetangga terdekat, misalkan ditetapkan K = 3.
2. Hitung jarak antara data baru dengan semua data training seperti pada tabel dibawah.

Tabel 2.3 Perhitungan Kuadrat Jarak Data Training Dengan Data Testing.

X1	X2	Kuadrat Jarak dengan Data Baru (3,7)
7	7	$(7-3)^2 + (7-7)^2 = 16$
7	4	$(7-3)^2 + (4-7)^2 = 25$
3	4	$(3-3)^2 + (4-7)^2 = 9$
1	4	$(1-3)^2 + (4-7)^2 = 13$

3. Urutkan hasil kuadrat tersebut secara ascending dan tetapkan tetangga terdekat berdasarkan nilai K (dimana $K = 3$) seperti pada Tabel

Tabel 2.4 Penentuan 3 Tetangga Terdekat Dari Data Testing

X1	X2	Euclidean	Peringkat Jarak	Tetangga Terdekat
7	7	$\sqrt{16} = 4$	3	YA
7	4	$\sqrt{25} = 5$	4	TIDAK
3	4	$\sqrt{9} = 3$	1	YA
1	4	$\sqrt{13} = 3,6$	2	YA

4. Berdasarkan hasil dari Tabel dimana nilai $K = 3$, mayoritas hasil Y yang diperoleh yaitu “BAIK”. Jadi hasil klasifikasi data testing dengan nilai $X1 = 3$ dan $X2 = 7$ adalah $Y = \text{Kelas Baik}$.

2.2.7 Root Mean Square Error (RMSE)

Dalam melihat hasil suatu prediksi dilakukan suatu evaluasi dimana evaluasi tersebut digunakan untuk mengetahui keakuratan hasil peramalan yang telah dilakukan terhadap data yang sebenarnya. Beberapa metode dapat digunakan untuk melakukan perhitungan kesalahan peramalan. Beberapa metode yang digunakan diantaranya adalah Root Mean Squared Error (RMSE). RMSE merupakan metode yang cukup sering digunakan dalam mengevaluasi kinerja prediksi. Dengan menggunakan RMSE, error yang ada menunjukkan seberapa besar perbedaan hasil estimasi dengan hasil yang akan diestimasi. Hal yang membuat berbeda karena adanya keacakan pada data atau karena tidak mengandung estimasi yang lebih akurat. RMSE merupakan mengakarkan nilai dari MSE yang sudah dicari sebelumnya. RMSE digunakan untuk mencari nilai keakurasian hasil peramalan dengan data history dengan menggunakan rumus yang ditunjukkan dengan nilai kesalahan. Semakin kecil nilai yang dihasilkan semakin bagus pula hasil peramalan yang dilakukan. Rumus RMSE dipresentasikan pada persamaan berikut[18]:

$$RMSE = \sqrt{\frac{\sum (y_t - \hat{y}_t)^2}{n}}$$

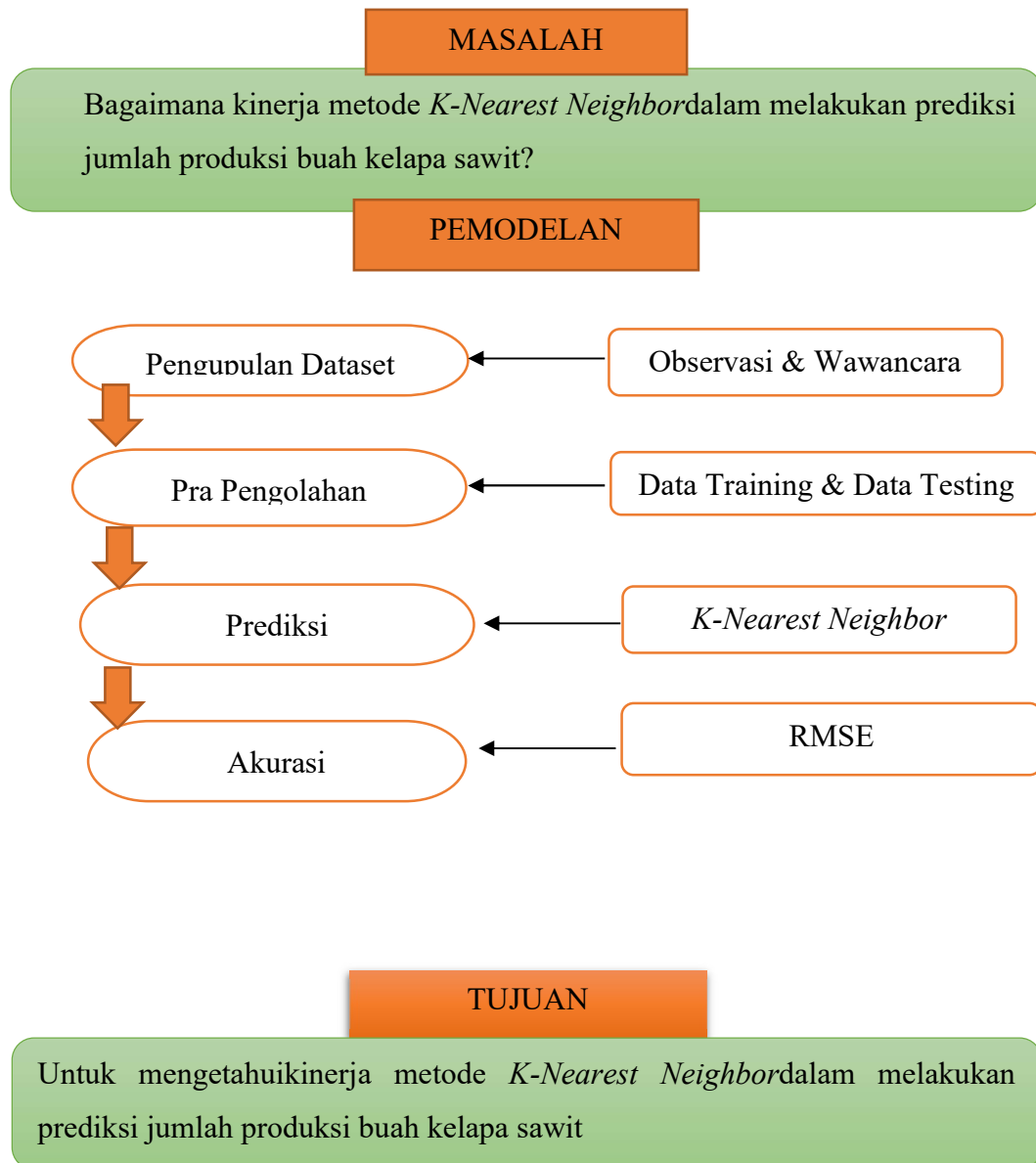
Keterangan :

Y_t =Nilai aktual indeks

\hat{Y}_t =Nilai prediksi indeks

N =Jumlah sampel

2.3 Kerangka Pikir



Gambar 2.6 Kerangka Pikir

BAB III

METODE PENELITIAN

3.1 Jenis, Metode, Subjek, objek, Waktu, dan Lokasi Penelitian

Dipandang dari jenis informasi yang akan diolah, maka penelitian ini merupakan penelitian Kuantitatif. Melihat tujuan dalam jenis penelitian ini adalah mengembangkan dan menggunakan model-model matematis, teori-teori dan/atau hipotesis yang berkaitan dengan fenomena alam

Metode yang digunakan dalam penelitian ini adalah metode penelitian eksperimen, dimana langkah tahapan dalam penelitian seperti. Pengumpulan data, Pengolahan Data, Prediksi menggunakan metode *K-Nearest Neighbor*, dan Evaluasi menggunakan Root Mean Square Error (RMSE) untuk mengetahui persentase penyimpangan hasil ramalan. Dimana yang menjadi subjek dan objek penelitian ini adalah Prediksi Jumlah Produksi Buah Kelapa Sawit Menggunakan Metode *K-Nearest Neighbor* Waktu yang dilakukan dalam melakukan pengumpulan data yakni 2 minggu yang dilakukan melalui tahapan observasi maupun wawancara. Dan lokasi penelitian atau Study kasus pada penelitian ini berada pada perusahaan PT. Hardaya Inti Plantations di Kabupaten Buol Provinsi Sulawesi Tengah.

3.2 Pengumpulan Data

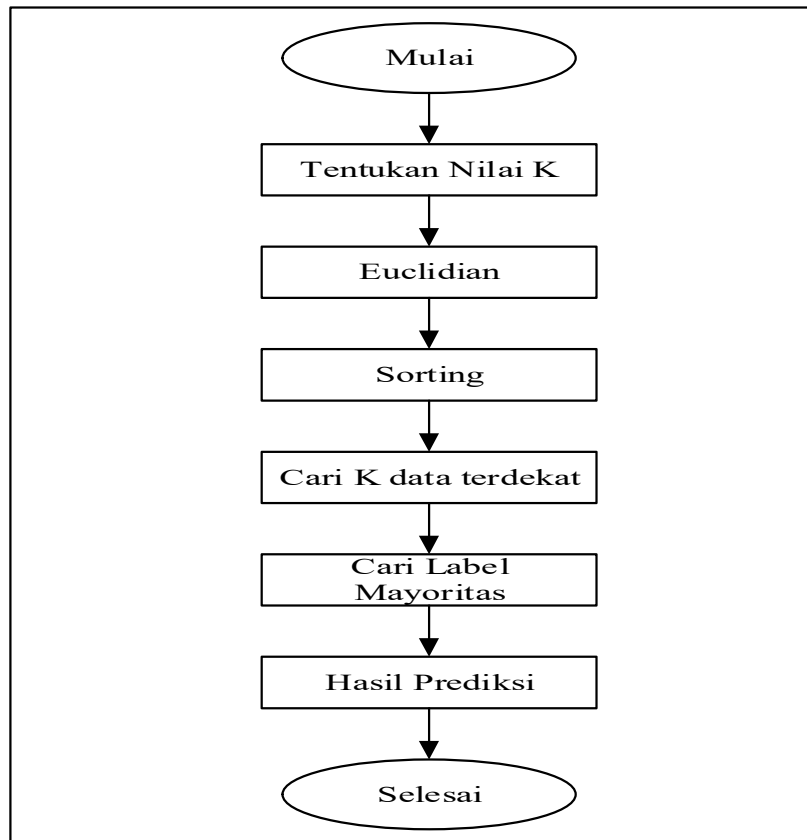
Sumber data yang digunakan dalam penelitian ini adalah data primer yaitu data yang di dapatkan secara langsung melalui teknik wawancara dengan Mill Manager PMKS sebagai pihak responden PT. Hardaya Inti Plantations mengenai data Jumlah Produksi Buah Kelapa Sawit tahun 2016-2021 terakhir dan kriteria penilaian yang akan digunakan dalam penelitian mengenai Prediksi jumlah produksi buah kelapa sawit. Sedangkan data sekunder diperoleh dengan mengumpulkan data atau keterangan melalui berbagai macam referensi yang ditulis oleh para ahli yang berhubungan dengan Produksi kelapa sawit maupun penelitian-penelitian yang membahas tentang Prediksi Jumlah Produksi Buah Kelapa Sawit Menggunakan Metode *K-Nearest Neighbor*. Baik yang bersumber dari buku ataupun jurnal terkait dari internet dan perpustakaan program studi Informatika pada Universitas Ichsan Gorontalo.

Adapun atribut data yang dilakukan pada penelitian ini yang membahas prediksi jumlah produksi kelapa sawit menggunakan metode *K-Nearest Neighbor* pada tabel berikut :

Tabel 3.1 Atribut data pada produksi buah kelapa sawit

No.	Produksi	Jumlah Produksi	Keterangan
1.	Januari	3.551.130	Banyak Produksi
2.	Februari	3.098.590	Banyak Produksi
3.	Maret	2.905.590	Banyak Produksi
4.	April	2.618.320	Banyak Produksi
5.	Mei	2.268.110	Banyak Produksi
6.	Juni	1.825.460	Sedikit Produksi
7.	Juli	1.668.800	Sedikit Produksi
8.	Agustus	2.778.820	Banyak Produksi
9.	September	6.439.000	Banyak Produksi
10.	Oktober	6.474.030	Banyak Produksi
11.	November	6.951.690	Banyak Produksi
12.	Desember	5.893.380	Banyak Produksi

3.3 Pemodelan



Gambar 3.1 *Model K-Nearest Neighbor*

3.3.1 Perhitungan *K-Nearest Neighbor*

Tujuan utama dilakukannya perhitungan *K-Nearest Neighbor* adalah untuk menentukan sebuah nilai $K = f(x)$ dengan nilai *deviasi* yang besar pada rentang ε sehingga dapat memprediksi nilai aktual untuk semua data latih. Sebuah parameter yang mewakili radius disekitar fungsi regresi disebut Variabel ε . SVR tidak mempermasalahkan adanya sebuah kesalahan jika kesalahan tersebut $< \varepsilon$.

Langkah-langkah untuk menghitung *Metode Algoritma K-Nearest Neighbor*:

1. Menentukan parameter K (jumlah tetangga paling dekat).
2. Menghitung kuadrat jarak *Euclid (query instance)* masing-masing objek terhadap data sampel yang diberikan.

3. Kemudian mengurutkan objek-objek tersebut kedalam kelompok yang mempunyai jarak *Euclid* terkecil.
4. Mengumpulkan kategori Y (*Klasifikasi K-Neares Neighbor*).
5. Dengan menggunakan kategori *Nearest Neighbor* yang paling mayoritas, maka dapat diprediksikan nilai *query instance* yang telah dihitung

3.3.2 Hasil Prediksi

Hasil prediksi merupakan hasil *output* pada data yang didapatkan dalam proses prediksi yang dilakukan oleh algoritma K-NN berdasarkan perhitungan yang telah dilakukan.

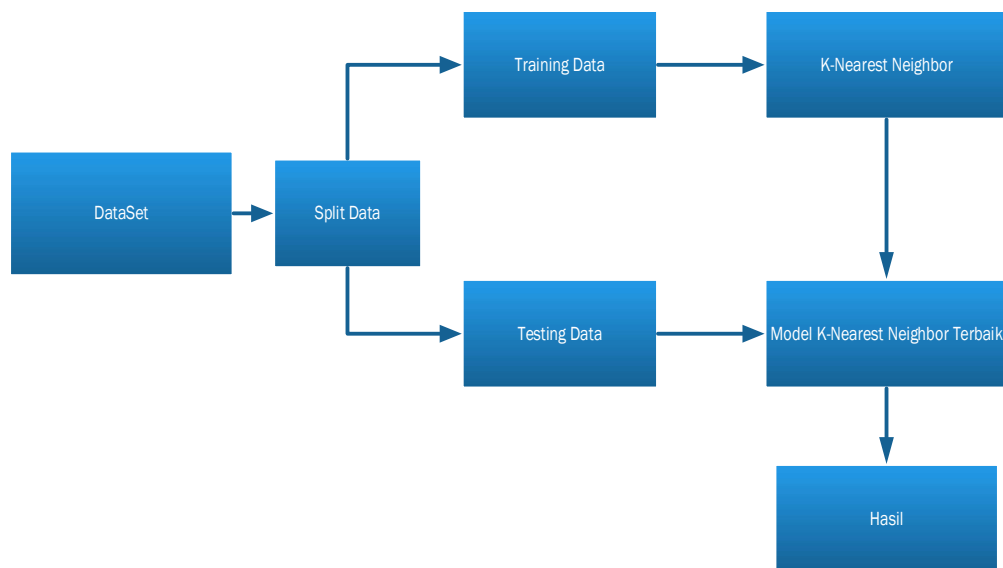
3.3.3 Akurasi

Model yang telah dihasilkan kemudian dilakukan uji akurasi dengan menggunakan *Root Mean Square Error (RMSE)* guna mengetahui persentase penyimpangan atau nilai *error* pada hasil prediksi jumlah produksi buah kelapa sawit .

BAB IV HASIL PENELITIAN

4.1. Model Usulan Penelitian

Model Usulan Pada Penelitian ini dapat digambar pada bagan alir sebagai berikut:



Gambar 4.1 Bagan Alir Model Usulan

Berdasarkan Gambar di atas dapat dijelaskan sistem kerja dari model yang diusulkan adalah dari dataset yang diperoleh dilakukan pembagian data menjadi dua yaitu data pemodelan dan data *hold out*. Data Pemodelan yang berisi data training digunakan untuk mencari model terbaik dari K-NN, Setelah model Terbaik didapatkan maka dilakukan pengujian akurasi dengan menggunakan data *hold out* yang berisi data *testing*.

Adapun Teknik pembagian data akan dijelaskan pada pembahasan berikutnya.

4.2. Dataset Penelitian

Dari Hasil Pengumpulan Data yang didapatkan pada PT. Hardaya Inti Plantations untuk 6 tahun terakhir yang selanjutnya akan digunakan sebagai *Dataset Penelitian*.

Dengan variabel yang digunakan adalah :

$X1 = \text{Luas Lahan (Ha)}$

$X2 = \text{Jumlah Pokok}$

$X3 = \text{Curah Hujan (mm)}$

$Y = \text{Produksi}$

Tabel 4.1 Hasil Pengumpulan Data

NO	TAHUN	BULAN	LUAS LAHAN (Ha)	JUMLAH POKOK	CURAH HUJAN (Mm)	PRODUKSI (Ton)
1	2016	JANUARI	3.087,15	335,797	148,2	3.551,130
2	2016	FEBRUARI	3.087,15	335,797	161,4	3.098,590
3	2016	MARET	3.087,15	335,797	250,6	2.905,590
4	2016	APRIL	3.087,15	335,797	109,7	2.618,320
5	2016	MEI	3.087,15	335,797	283,6	2.268,110
6	2016	JUNI	3.087,15	335,797	320,0	1.825,460
7	2016	JULI	3.087,15	335,797	195,1	1.668,800
8	2016	AGUSTUS	3.087,15	335,797	194,6	2.778,820
9	2016	SEPTEMBER	3.087,15	335,797	264,2	6.439,000
10	2016	OKTOBER	3.087,15	335,797	330,2	6.474,030
11	2016	NOVEMBER	3.087,15	335,797	243,5	6.951,690
12	2016	DESEMBER	3.087,15	335,797	146,4	5.893,380
13	2017	JANUARI	3.078,02	329,042	259,1	5.025,150
14	2017	FEBRUARI	3.078,02	329,042	177,3	3.861,600
15	2017	MARET	3.078,02	329,042	238,2	4.034,010
16	2017	APRIL	3.078,02	329,042	259,4	4.529,840
17	2017	MEI	3.078,02	329,042	305,0	4.681,830
18	2017	JUNI	3.078,02	329,042	142,2	3.629,540
19	2017	JULI	3.078,02	329,042	137,1	5.316,800
20	2017	AGUSTUS	3.078,02	329,042	216,7	4.541,430
21	2017	SEPTEMBER	3.078,02	329,042	155,6	4.425,880
22	2017	OKTOBER	3.078,02	329,042	326,4	4.156,600

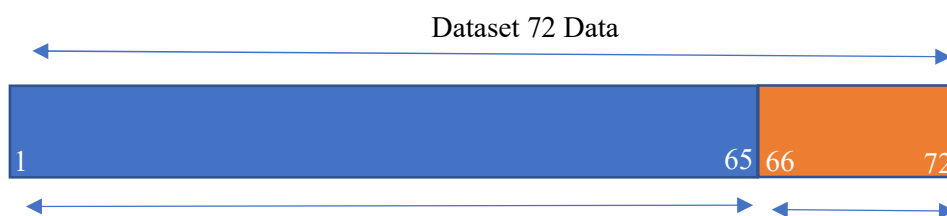
NO	TAHUN	BULAN	LUAS LAHAN (Ha)	JUMLAH POKOK	CURAH HUJAN (Mm)	PRODUKSI (Ton)
23	2017	NOVEMBER	3.078,02	329,042	297,0	3.991,600
24	2017	DESEMBER	3.078,02	329,042	210,5	2.886,216
25	2018	JANUARI	3.113,61	338,649	184,7	2.324,300
26	2018	FEBRUARI	3.113,61	338,649	130,4	1.827,540
27	2018	MARET	3.113,61	338,649	160,2	2.385,655
28	2018	APRIL	3.113,61	338,649	209,0	3.281,659
29	2018	MEI	3.113,61	338,649	158,1	4.366,220
30	2018	JUNI	3.113,61	338,649	162,4	4.308,350
31	2018	JULI	3.113,61	338,649	129,6	7.031,950
32	2018	AGUSTUS	3.113,61	338,649	284,0	6.560,950
33	2018	SEPTEMBER	3.113,61	338,649	272,1	5.533,410
34	2018	OKTOBER	3.113,61	338,649	354,0	6.179,635
35	2018	NOVEMBER	3.113,61	338,649	362,6	5.393,710
36	2018	DESEMBER	3.113,61	338,649	332,4	4.671,290
37	2019	JANUARI	3.098,20	330,579	220,4	4.401,800
38	2019	FEBRUARI	3.098,20	330,579	105,2	3.133,540
39	2019	MARET	3.098,20	330,579	168,1	3.494,660
40	2019	APRIL	3.098,20	330,579	147,2	2.745,090
41	2019	MEI	3.098,20	330,579	200,0	2.456,013
42	2019	JUNI	3.098,20	330,579	134,4	2.208,365
43	2019	JULI	3.098,20	330,579	204,5	3.562,354
44	2019	AGUSTUS	3.098,20	330,579	163,2	3.968,660
45	2019	SEPTEMBER	3.098,20	330,579	301,0	4.597,630
46	2019	OKTOBER	3.098,20	330,579	189,3	5.912,060
47	2019	NOVEMBER	3.098,20	330,579	190,0	6.075,710
48	2019	DESEMBER	3.098,20	330,579	172,2	4.918,360
49	2020	JANUARI	3.127,74	314,104	202,6	3.874,180
50	2020	FEBRUARI	3.127,74	314,104	125,2	4.207,790
51	2020	MARET	3.127,74	314,104	186,4	3.931,300
52	2020	APRIL	3.127,74	314,104	145,2	4.326,110
53	2020	MEI	3.127,74	314,104	268,6	3.369,090
54	2020	JUNI	3.127,74	314,104	110,8	3.927,060
55	2020	JULI	3.127,74	314,104	221,1	2.264,960
56	2020	AGUSTUS	3.127,74	314,104	320,7	2.981,940
57	2020	SEPTEMBER	3.127,74	314,104	146,6	3.579,930
58	2020	OKTOBER	3.127,74	314,104	301,3	4.237,770
59	2020	NOVEMBER	3.127,74	314,104	138,3	4.772,620
60	2020	DESEMBER	3.127,74	314,104	128,6	4.656,060

NO	TAHUN	BULAN	LUAS LAHAN (Ha)	JUMLAH POKOK	CURAH HUJAN (Mm)	PRODUKSI (Ton)
61	2021	JANUARI	3.129,50	337,305	313,2	4.417,820
62	2021	FEBRUARI	3.129,50	337,305	152,4	3.745,230
63	2021	MARET	3.129,50	337,305	168,5	3.946,400
64	2021	APRIL	3.129,50	337,305	142,2	3.846,040
65	2021	MEI	3.129,50	337,305	218,6	2.278,360
66	2021	JUNI	3.129,50	337,305	115,7	4.020,790
67	2021	JULI	3.129,50	337,305	221,0	3.075,660
68	2021	AGUSTUS	3.129,50	337,305	340,7	2.772,850
69	2021	SEPTEMBER	3.129,50	337,305	164,6	3.224,432
70	2021	OKTOBER	3.129,50	337,305	300,2	3.535,615
71	2021	NOVEMBER	3.129,50	337,305	221,1	3.812,423
72	2021	DESEMBER	3.129,50	337,305	173,3	3.639,928

(Sumber PT. Hardaya Inti Plantations 2021)

4.3. Pembagian Data

Pembagian Data dilakukan dengan membagi 2 data yaitu data pemodelan dan data *hold out*. Dari Total 72 Record Dataset hasil pembagian didapatkan data pemodelan sebanyak 90% (65 data) dan Data *hold Out* 10% (7 Data).



Gambar 4.2 Pembagian Data Training 72 Data

Dari hasil Pembagian data tersebut. Data dengan record 1 s/d 65 akan dijadikan data training untuk membangun model K-NN

Tabel 4.2 Data Training

<i>NO</i>	<i>X1</i>	<i>X2</i>	<i>X3</i>	<i>y</i>
1	3.087,15	335,797	148,2	3.551,130
2	3.087,15	335,797	161,4	3.098,590
3	3.087,15	335,797	250,6	2.905,590
4	3.087,15	335,797	109,7	2.618,320
5	3.087,15	335,797	283,6	2.268,110
6	3.087,15	335,797	320,0	1.825,460
7	3.087,15	335,797	195,1	1.668,800
8	3.087,15	335,797	194,6	2.778,820
9	3.087,15	335,797	264,2	6.439,000
10	3.087,15	335,797	330,2	6.474,030
11	3.087,15	335,797	243,5	6.951,690
12	3.087,15	335,797	146,4	5.893,380
13	3.078,02	329,042	259,1	5.025,150
14	3.078,02	329,042	177,3	3.861,600
15	3.078,02	329,042	238,2	4.034,010
16	3.078,02	329,042	259,4	4.529,840
17	3.078,02	329,042	305,0	4.681,830
18	3.078,02	329,042	142,2	3.629,540
19	3.078,02	329,042	137,1	5.316,800
20	3.078,02	329,042	216,7	4.541,430
21	3.078,02	329,042	155,6	4.425,880
22	3.078,02	329,042	326,4	4.156,600
23	3.078,02	329,042	297,0	3.991,600
24	3.078,02	329,042	210,5	2.886,216
25	3.113,61	338,649	184,7	2.324,300
26	3.113,61	338,649	130,4	1.827,540
27	3.113,61	338,649	160,2	2.385,655
28	3.113,61	338,649	209,0	3.281,659
29	3.113,61	338,649	158,1	4.366,220
30	3.113,61	338,649	162,4	4.308,350
31	3.113,61	338,649	129,6	7.031,950
32	3.113,61	338,649	284,0	6.560,950
33	3.113,61	338,649	272,1	5.533,410
34	3.113,61	338,649	354,0	6.179,635
35	3.113,61	338,649	362,6	5.393,710
36	3.113,61	338,649	332,4	4.671,290
37	3.098,20	330,579	220,4	4.401,800

<i>NO</i>	<i>X1</i>	<i>X2</i>	<i>X3</i>	<i>y</i>
38	3.098,20	330,579	105,2	3.133,540
39	3.098,20	330,579	168,1	3.494,660
40	3.098,20	330,579	147,2	2.745,090
41	3.098,20	330,579	200,0	2.456,013
42	3.098,20	330,579	134,4	2.208,365
43	3.098,20	330,579	204,5	3.562,354
44	3.098,20	330,579	163,2	3.968,660
45	3.098,20	330,579	301,0	4.597,630
46	3.098,20	330,579	189,3	5.912,060
47	3.098,20	330,579	190,0	6.075,710
48	3.098,20	330,579	172,2	4.918,360
49	3.127,74	314,104	202,6	3.874,180
50	3.127,74	314,104	125,2	4.207,790
51	3.127,74	314,104	186,4	3.931,300
52	3.127,74	314,104	145,2	4.326,110
53	3.127,74	314,104	268,6	3.369,090
54	3.127,74	314,104	110,8	3.927,060
55	3.127,74	314,104	221,1	2.264,960
56	3.127,74	314,104	320,7	2.981,940
57	3.127,74	314,104	146,6	3.579,930
58	3.127,74	314,104	301,3	4.237,770
59	3.127,74	314,104	138,3	4.772,620
60	3.127,74	314,104	128,6	4.656,060
61	3.129,50	337,305	313,2	4.417,820
62	3.129,50	337,305	152,4	3.745,230
63	3.129,50	337,305	168,5	3.946,400
64	3.129,50	337,305	142,2	3.846,040
65	3.129,50	337,305	218,6	2.278,360

Selanjutnya 10% data terakhir yakni sebanyak 10 record akan dijadikan data testing untuk mengevaluasi model K-NN yang akan dibangun

Tabel 4.3 Data Testing

<i>NO</i>	<i>X1</i>	<i>X2</i>	<i>X3</i>	<i>y</i>
66	3.129,50	337,305	115,7	4.020,790
67	3.129,50	337,305	221,0	3.075,660
68	3.129,50	337,305	340,7	2.772,850
69	3.129,50	337,305	164,6	3.224,432
70	3.129,50	337,305	300,2	3.535,615
71	3.129,50	337,305	221,1	3.812,423
72	3.129,50	337,305	173,3	3.639,928

4.4. Pembentukan Model K-NN

Berikut Merupakan Tahapan Dari Metode K-Nearest Neighbor (K-NN)[6]

1. Tentukan parameter K = Jumlah dari Tetangga Terdekat (*nearest neighbor*)
2. Hitung jarak data baru yang dinyatakan dengan seluruh sampel data *training* (pelatihan) dengan menggunakan Rumus Euclidean Distance
3. Urutkan seluruh jarak berdasarkan jarak *minimum* dan tetapkan Tetangga Terdekat sesuai dengan nilai K
4. Sesuaikan klasifikasi dari Y dengan persekitaran yang telah ditetapkan
5. Gunakan kelas dengan jumlah terbanyak sebagai dasar menentukan kelas dari data baru yang dinyatakan.

Berdasarkan Dataset Yang ada maka rumus euclidean Distance yang digunakan pada penelitian ini adalah:

$$Jarak = \sqrt{(x1tr - x1ts)^2 + (x2r - x2ts)^2 + (x3tr - x3ts)^2}$$

4.5. Perhitungan K-NN

Berikut ini merupakan contoh dari penerapan metode K-Nearest Neighbor menggunakan rumus Euclidean Distance

Langkah 1 : Menghitung Jarak Antara Data Training dan Data Testing

Tabel 4.4 Data Training

<i>No</i>	<i>X1</i>	<i>X2</i>	<i>X3</i>	<i>Y</i>
1	3.087,15	335,797	148,2	3.551,130
2	3.087,15	335,797	160,4	3.098,590
3	3.087,15	335,797	230,6	2.905,590
4	3.087,15	335,797	109,7	2.618,320
5	3.087,15	335,797	283,6	2.268,110
6	3.087,15	335,797	320,0	1.825,460
7	3.087,15	335,797	195,1	1.668,800
8	3.087,15	335,797	194,6	2.778,820
9	3.087,15	335,797	264,2	6.439,000
10	3.087,15	335,797	330,2	6.474,030
..				
65	3.129,50	337,305	218,6	2.278,360

Tabel 4.5 Data Testing

<i>No</i>	<i>X1</i>	<i>X2</i>	<i>X3</i>	<i>Y</i>
66	3.129,50	337,305	115,7	?

Pada Data Testing nilai Y akan diprediksi Kembali untuk mengetahui perbedaan antara data asli(Aktual) dengan data Prediksi (Y')

$$\begin{aligned}
 \text{Jarak1} &= \sqrt{(x1tr - X1ts)^2 + (x2tr - x2ts)^2 + (x3tr - x3ts)^2} \\
 &= \sqrt{(3.087,15 - 3.129,50)^2 + (335,797 - 337,305)^2 + (148,2 - 115,7)^2} \\
 &= \sqrt{1793,552 + 2,274 + 1056,25}
 \end{aligned}$$

$$\sqrt{2852,047}$$

$$= 53,404$$

$$\begin{aligned} \text{Jarak2} &= \sqrt{(x1tr - X1ts)^2 + (x2tr - x2ts)^2 + (x3tr - x3ts)^2} \\ &= \sqrt{(3.087,15 - 3.129,50)^2 + (335,797 - 337,305)^2 + (161,4 - 115,7)^2} \\ &= \sqrt{1793,552 + 2,274 + 2088,49} \\ &= \sqrt{3884,287} \\ &= 62,324 \end{aligned}$$

$$\begin{aligned} \text{Jarak3} &= \sqrt{(x1tr - X1ts)^2 + (x2tr - x2ts)^2 + (x3tr - x3ts)^2} \\ &= \sqrt{(3.087,15 - 3.129,50)^2 + (335,797 - 337,305)^2 + (250,6 - 115,7)^2} \\ &= \sqrt{1793,552 + 2,274 + 141,399} \\ &= \sqrt{19993,81} \\ &= 141,399 \end{aligned}$$

..

Perhitungan jarak dilakukan seterusnya sampai data ke 65 Sehingga didapatkan tabel jarak sebagai berikut

Tabel 4.6 Hasil Perhitungan Jarak

No.	X1	X2	X3	Y	Jarak
1	1793,522	2,274064	1056,25	3551,13	53,40456
2	1793,522	2,274064	2088,49	3098,59	62,32404
3	1793,522	2,274064	18198,01	2905,59	141,3995
4	1793,522	2,274064	36	2618,32	42,79949
5	1793,522	2,274064	28190,41	2268,11	173,1653
6	1793,522	2,274064	41738,49	1825,46	208,6487
7	1793,522	2,274064	6304,36	1668,8	90,00087
8	1793,522	2,274064	6225,21	2778,82	89,56007
9	1793,522	2,274064	22052,25	6439	154,4281
10	1793,522	2,274064	46010,25	6474,03	218,6459

<i>No.</i>	<i>X1</i>	<i>X2</i>	<i>X3</i>	<i>Y</i>	<i>Jarak</i>
11	1793,522	2,274064	16332,84	6951,69	134,6426
12	1793,522	2,274064	942,49	5893,38	52,32864
13	2650,19	68,27717	20563,56	5025,15	152,5845
14	2650,19	68,27717	3794,56	3861,6	80,70333
15	2650,19	68,27717	15006,25	4034,01	133,1342
16	2650,19	68,27717	20649,69	4529,84	152,8665
17	2650,19	68,27717	35834,49	4681,83	196,3491
18	2650,19	68,27717	702,25	3629,54	58,4869
19	2650,19	68,27717	457,96	5316,8	56,3598
20	2650,19	68,27717	10201	4541,43	113,6638
21	2650,19	68,27717	1592,01	4425,88	65,65423
22	2650,19	68,27717	44394,49	4156,6	217,0552
23	2650,19	68,27717	32869,69	3991,6	188,6482
24	2650,19	68,27717	8987,04	2886,216	108,192
25	252,4921	1,806336	4761	2324,3	70,81877
26	252,4921	1,806336	216,09	1827,54	21,68844
27	252,4921	1,806336	1980,25	2385,655	47,27101
28	252,4921	1,806336	8704,89	3281,659	94,65299
29	252,4921	1,806336	1797,76	4366,22	45,29965
30	252,4921	1,806336	2180,89	4308,35	49,34763
31	252,4921	1,806336	193,21	7031,95	21,1544
32	252,4921	1,806336	28324,89	6560,95	169,0538
33	252,4921	1,806336	24460,96	5533,41	157,2109
34	252,4921	1,806336	56786,89	6179,635	238,833
35	252,4921	1,806336	60959,61	5393,71	247,4144
36	252,4921	1,806336	46958,89	4671,29	217,286
37	979,69	45,23908	10962,09	4401,8	109,4852
38	979,69	45,23908	110,25	3133,54	33,69242
39	979,69	45,23908	2745,76	3494,66	61,40594
40	979,69	45,23908	992,25	2745,09	44,91302
41	979,69	45,23908	7106,49	2456,013	90,17438
42	979,69	45,23908	349,69	2208,365	37,07586
43	979,69	45,23908	7885,44	3562,354	94,39475
44	979,69	45,23908	2256,25	3968,66	57,28158
45	979,69	45,23908	34336,09	4597,63	188,0453
46	979,69	45,23908	5416,96	5912,06	80,26138
47	979,69	45,23908	5520,49	6075,71	80,90376
48	979,69	45,23908	3192,25	4918,36	64,93981
49	3,0976	538,2864	7551,61	3874,18	89,96107
50	3,0976	538,2864	90,25	4207,79	25,13233

No.	$X1$	$X2$	$X3$	Y	Jarak
51	3,0976	538,2864	4998,49	3931,3	74,43033
52	3,0976	538,2864	870,25	4326,11	37,57172
53	3,0976	538,2864	23378,41	3369,09	154,6603
54	3,0976	538,2864	24,01	3927,06	23,77802
55	3,0976	538,2864	11109,16	2264,96	107,9377
56	3,0976	538,2864	42025	2981,94	206,3162
57	3,0976	538,2864	954,81	3579,93	38,68067
58	3,0976	538,2864	34447,36	4237,77	187,0528
59	3,0976	538,2864	510,76	4772,62	32,43677
60	3,0976	538,2864	166,41	4656,06	26,6044
61	0	0	39006,25	4417,82	197,5
62	0	0	1346,89	3745,23	36,7
63	0	0	2787,84	3946,4	52,8
64	0	0	702,25	3846,04	26,5
65	0	0	10588,41	2278,36	102,9

Langkah 2. Menentukan Peringkat Jarak

Peringkat Jarak ditentukan dengan nilai jarak paling kecil sebagai peringkat pertama dan seterusnya:

Tabel 4.7 Peringkat Jarak

No.	$X1$	$X2$	$X3$	Y	Jarak	Peringkat
31	252,4921	1,806336	193,21	7031,95	21,1544	1
26	252,4921	1,806336	216,09	1827,54	21,68844	2
54	3,0976	538,2864	24,01	3927,06	23,77802	3
50	3,0976	538,2864	90,25	4207,79	25,13233	4
64	0	0	702,25	3846,04	26,5	5
60	3,0976	538,2864	166,41	4656,06	26,6044	6
59	3,0976	538,2864	510,76	4772,62	32,43677	7
38	979,69	45,23908	110,25	3133,54	33,69242	8
62	0	0	1346,89	3745,23	36,7	9
42	979,69	45,23908	349,69	2208,365	37,07586	10
52	3,0976	538,2864	870,25	4326,11	37,57172	11
57	3,0976	538,2864	954,81	3579,93	38,68067	12
4	1793,522	2,274064	36	2618,32	42,79949	13
40	979,69	45,23908	992,25	2745,09	44,91302	14
29	252,4921	1,806336	1797,76	4366,22	45,29965	15
27	252,4921	1,806336	1980,25	2385,655	47,27101	16

No.	X1	X2	X3	Y	Jarak	Peringkat
30	252,4921	1,806336	2180,89	4308,35	49,34763	17
12	1793,522	2,274064	942,49	5893,38	52,32864	18
63	0	0	2787,84	3946,4	52,8	19
1	1793,522	2,274064	1056,25	3551,13	53,40456	20
19	2650,19	68,27717	457,96	5316,8	56,3598	21
44	979,69	45,23908	2256,25	3968,66	57,28158	22
18	2650,19	68,27717	702,25	3629,54	58,4869	23
39	979,69	45,23908	2745,76	3494,66	61,40594	24
2	1793,522	2,274064	2088,49	3098,59	62,32404	25
48	979,69	45,23908	3192,25	4918,36	64,93981	26
21	2650,19	68,27717	1592,01	4425,88	65,65423	27
25	252,4921	1,806336	4761	2324,3	70,81877	28
51	3,0976	538,2864	4998,49	3931,3	74,43033	29
46	979,69	45,23908	5416,96	5912,06	80,26138	30
14	2650,19	68,27717	3794,56	3861,6	80,70333	31
47	979,69	45,23908	5520,49	6075,71	80,90376	32
8	1793,522	2,274064	6225,21	2778,82	89,56007	33
49	3,0976	538,2864	7551,61	3874,18	89,96107	34
7	1793,522	2,274064	6304,36	1668,8	90,00087	35
41	979,69	45,23908	7106,49	2456,013	90,17438	36
43	979,69	45,23908	7885,44	3562,354	94,39475	37
28	252,4921	1,806336	8704,89	3281,659	94,65299	38
65	0	0	10588,41	2278,36	102,9	39
55	3,0976	538,2864	11109,16	2264,96	107,9377	40
24	2650,19	68,27717	8987,04	2886,216	108,192	41
37	979,69	45,23908	10962,09	4401,8	109,4852	42
20	2650,19	68,27717	10201	4541,43	113,6638	43
15	2650,19	68,27717	15006,25	4034,01	133,1342	44
11	1793,522	2,274064	16332,84	6951,69	134,6426	45
3	1793,522	2,274064	18198,01	2905,59	141,3995	46
13	2650,19	68,27717	20563,56	5025,15	152,5845	47
16	2650,19	68,27717	20649,69	4529,84	152,8665	48
9	1793,522	2,274064	22052,25	6439	154,4281	49
53	3,0976	538,2864	23378,41	3369,09	154,6603	50
33	252,4921	1,806336	24460,96	5533,41	157,2109	51
32	252,4921	1,806336	28324,89	6560,95	169,0538	52
5	1793,522	2,274064	28190,41	2268,11	173,1653	53
58	3,0976	538,2864	34447,36	4237,77	187,0528	54
45	979,69	45,23908	34336,09	4597,63	188,0453	55
23	2650,19	68,27717	32869,69	3991,6	188,6482	56
17	2650,19	68,27717	35834,49	4681,83	196,3491	57

No.	X1	X2	X3	Y	Jarak	Peringkat
61	0	0	39006,25	4417,82	197,5	58
56	3,0976	538,2864	42025	2981,94	206,3162	59
6	1793,522	2,274064	41738,49	1825,46	208,6487	60
22	2650,19	68,27717	44394,49	4156,6	217,0552	61
36	252,4921	1,806336	46958,89	4671,29	217,286	62
10	1793,522	2,274064	46010,25	6474,03	218,6459	63
34	252,4921	1,806336	56786,89	6179,635	238,833	64
35	252,4921	1,806336	60959,61	5393,71	247,4144	65

Langkah 3. Menentukan Hasil Prediksi Berdasarkan Nilai K

a). Uji Coba Dengan K=3

Untuk Uji Coba Pertama akan digunakan 3 data dengan peringkat teratas berdasarkan nilai K=3

Tabel 4.8 Nilai K= 3

No.	X1	X2	X3	Y	Jarak	Peringkat
31	252,4921	1,806336	193,21	7031,95	21,1544	1
26	252,4921	1,806336	216,09	1827,54	21,68844	2
54	3,0976	538,2864	24,01	3927,06	23,77802	3

Maka Hasil Prediksi Berdasarkan data testing adalah:

$$\begin{aligned}
 Y &= \frac{y_1 + y_2 + y_3}{k} \\
 &= \frac{7031,95 + 1827,54 + 3927,06}{3} \\
 &= \frac{12786,55}{3} = 4262,183
 \end{aligned}$$

Berdasarkan Hasil Perhitungan tersebut didapatkan hasil Prediksi sebesar 4262,183 Ton

b). Uji Coba Dengan K=5

Untuk Uji Coba Pertama akan digunakan 5 data dengan peringkat teratas berdasarkan nilai K=5

Tabel 4.9 Nilai K= 5

No.	X1	X2	X3	Y	Jarak	Peringkat
31	252,4921	1,806336	193,21	7031,95	21,1544	1
26	252,4921	1,806336	216,09	1827,54	21,68844	2
54	3,0976	538,2864	24,01	3927,06	23,77802	3
50	3,0976	538,2864	90,25	4207,79	25,13233	4
64	0	0	702,25	3846,04	26,5	5

Maka Hasil Prediksi Berdasarkan data testing adalah:

$$\begin{aligned}
 Y &= \frac{y_1 + y_2 + y_3 + y_4 + y_5}{k} \\
 &= \frac{7031,95 + 1827,54 + 3927,06 + 4207,79 + 3846,04}{5} \\
 &= \frac{20840,38}{5} = 4168,076
 \end{aligned}$$

Berdasarkan Hasil Perhitungan tersebut didapatkan hasil Prediksi sebesar 4168,076 Ton

C. Uji Coba Dengan K=7

Untuk Uji Coba Pertama akan digunakan 7 data dengan peringkat teratas berdasarkan nilai K=7

Tabel 4.10 Nilai K= 7

No.	X1	X2	X3	Y	Jarak	Peringkat
31	252,4921	1,806336	193,21	7031,95	21,1544	1
26	252,4921	1,806336	216,09	1827,54	21,68844	2
54	3,0976	538,2864	24,01	3927,06	23,77802	3
50	3,0976	538,2864	90,25	4207,79	25,13233	4
64	0	0	702,25	3846,04	26,5	5
60	3,0976	538,2864	166,41	4656,06	26,6044	6
59	3,0976	538,2864	510,76	4772,62	32,43677	7

Maka Hasil Prediksi Berdasarkan data testing adalah:

$$Y = \frac{y_1 + y_2 + y_3 + y_4 + y_5 + y_6 + y_7}{k}$$

$$= \frac{7031,95 + 1827,54 + 3927,06 + 420,79 + 3846,04 + 26,604 + 32,43}{7}$$

$$= \frac{30269,06}{7} = 4324,151$$

Berdasarkan Hasil Perhitungan tersebut didapatkan hasil Prediksi sebesar 4324,151Ton

Dari 3 Kali Percobaan yang dilakukan untuk masing masing nilai K=3, 5 dan 7 Terhadap Data Testing No. 66 maka secara rinci dapat dilihat pada tabel berikut

Tabel 4.11 Perbandingan Data Aktual dengan Data Pediksi Untuk Masing-Masing Nilai K

NO	X1	X2	X3	Y Aktual	Y'Prediksi (K=3)	Y'Prediksi (K=5)	Y'Prediksi (K=7)
66	3.129,50	337,305	115,7	4.020,790	4.262,183	4.168,076	4.324,151

Dari Tabel tersebut dapat ditunjukkan bahwa perbedaan antara data aktual dan data prediksi tidak memperlihatkan perbedaan yang signifikan.

4.6 Hasil Prediksi

Pada Tabel berikut, terlihat bahwa ada dua warna yang berbeda antara hijau dan kuning. Hijau menunjukkan data sebenarnya, sedangkan kuning adalah data nilai prediksi yang dihasilkan

Tabel 4.12. Hasil Prediksi dengan K=3

NO	X1	X2	X3	y	y'
66	3.129,50	337,305	115,7	4.020,790	4.262,183

<i>NO</i>	<i>X1</i>	<i>X2</i>	<i>X3</i>	<i>y</i>	<i>y'</i>
67	3.129,50	337,305	221,0	3.075,660	2.608,326
68	3.129,50	337,305	340,7	2.772,850	5414,878
69	3.129,50	337,305	164,6	3.224,432	3.999,933
70	3.129,50	337,305	300,2	3.535,615	3.882,065
71	3.129,50	337,305	221,1	3.812,423	3.214,645
72	3.129,50	337,305	173,3	3.639,928	3.443,356

Tabel 4.13. Hasil Prediksi dengan K=5

<i>NO</i>	<i>X1</i>	<i>X2</i>	<i>X3</i>	<i>y</i>	<i>y'</i>
66	3.129,50	337,305	115,7	4.020,790	4.168,006
67	3.129,50	337,305	221,0	3.075,660	3.283,104
68	3.129,50	337,305	340,7	2.772,850	3.554,183
69	3.129,50	337,305	164,6	3.224,432	3.432,465
70	3.129,50	337,305	300,2	3.535,615	3.882,065
71	3.129,50	337,305	221,1	3.812,423	3.877,545
72	3.129,50	337,305	173,3	3.639,928	3.674,813

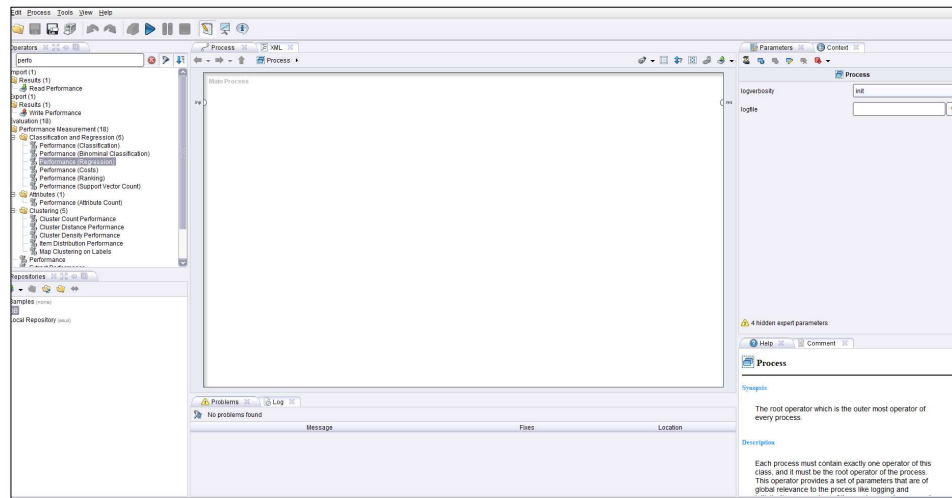
Tabel 4.14. Hasil Prediksi dengan K=7

<i>NO</i>	<i>X1</i>	<i>X2</i>	<i>X3</i>	<i>y</i>	<i>y'</i>
66	3.129,50	337,305	115,7	4.020,790	4.324,151
67	3.129,50	337,305	221,0	3.075,660	3.104,224
68	3.129,50	337,305	340,7	2.772,850	3.452,117
69	3.129,50	337,305	164,6	3.224,432	3.412,332
70	3.129,50	337,305	300,2	3.535,615	3.896,213
71	3.129,50	337,305	221,1	3.812,423	3.683,712
72	3.129,50	337,305	173,3	3.639,928	3.514,453

BAB V HASIL DAN PEMBAHASAN

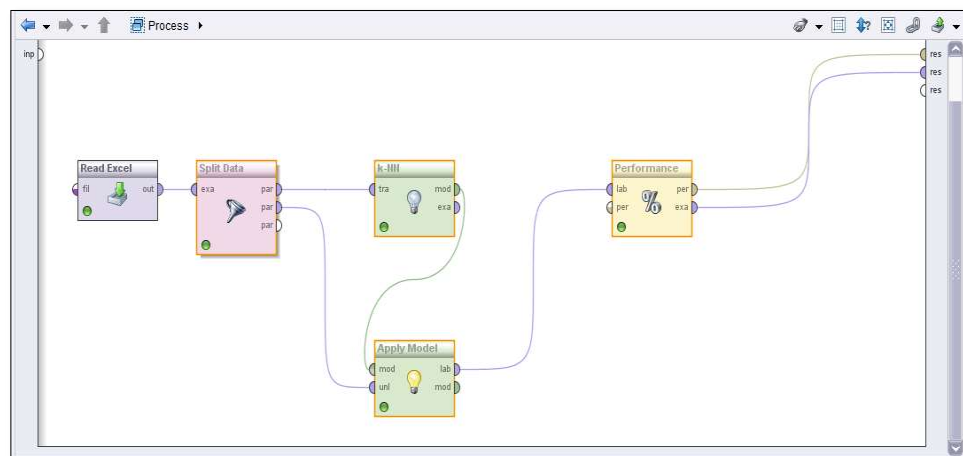
5.1. Hasil Pemodelan Dengan Rapid Miner

5.1.1. Tahapan Pemodelan Pada Rapid Miner



Gambar 5.1 Memulai Rapid Miner

Pada Penelitian Ini Model K-NN diuji Coba dengan Menggunakan Aplikasi *Ripied Miner*. Gambar 5.1 adalah Halaman Pada Saat Aplikasi Rapid Miner pertama kali dijalankan



Gambar 5.2 Pemodelan K-NN Pada Rapid Miner

Secara Lengkap Pemodelan Knn untuk prediksi Produksi Kelapa sawit dapat dilihat pada gambar 5.2. dimana dataset yang digunakan dalam bentuk excel

5.1.1.1. Pengaturan Atribut Variabel

Data import wizard - Step 4 of 4

This wizard guides you to import your data.
Step 4: RapidMiner uses strongly typed attributes. In this step, you can define the data types of your attributes. Furthermore, RapidMiner assigns roles to the attributes, defining what they can be used for by the individual operators. These roles can be also defined here. Finally, you can rename attributes or deselect them entirely.

☒ Preview uses only first 100 rows.

<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
TAHUN	BULAN	LUAS LAHA	JUMLAH PO	CURAH HU.	PRODUKSI
integer	polyno...	real	real	numeric	numeric
attribute	attribute	attribute	attribute	attribute	attribute
2016	JANUARI	3087.150	335.797	148.200	3551.130
2016	FEBRUARI	3087.150	335.797	161.400	3098.590
2016	MARET	3087.150	335.797	250.600	2905.590
2016	APRIL	3087.150	335.797	109.700	2618.320
2016	MEI	3087.150	335.797	283.600	2268.110
2016	JUNI	3087.150	335.797	320	1825.460
2016	JULI	3087.150	335.797	195.100	1668.800
2016	AGUSTUS	3087.150	335.797	194.600	2778.820
2016	SEPTEMBER	3087.150	335.797	264.200	6439
2016	OKTOBER	3087.150	335.797	330.200	6474.030
2016	NOVEMBER	3087.150	335.797	243.500	6951.690
2016	DESEMBER	3087.150	335.797	146.400	5893.380
2017	JANUARI	3078.020	329.042	259.100	5025.150
2017	FEBRUARI	3078.020	329.042	177.300	3861.600
2017	MARET	3078.020	329.042	238.200	4034.010
2017	APRIL	3078.020	329.042	259.400	4529.840
2017	MEI	3078.020	329.042	305	4681.830
2017	JUNI	3078.020	329.042	142.200	3629.540
2017	JULI	3078.020	329.042	137.100	5316.800

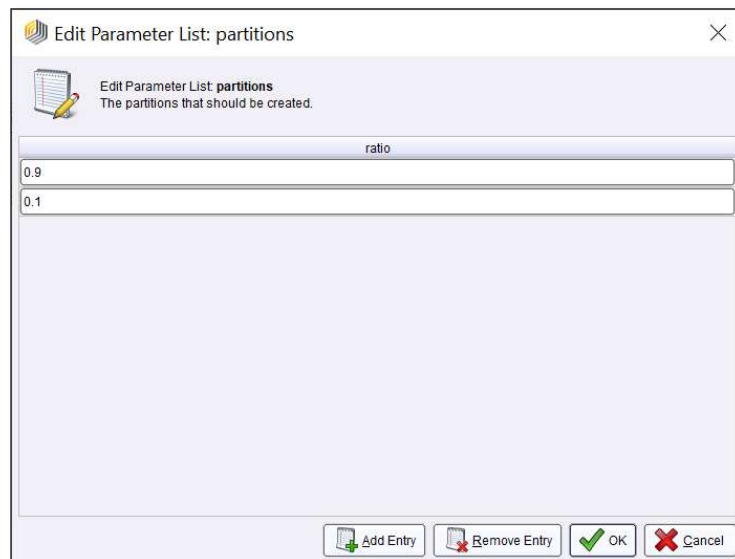
☒ 0 errors. ☒ Ignore errors ☐ Show only errors

Row, Column	Error	Original value	Message
-------------	-------	----------------	---------

Gambar 5.3 Pengaturan Atribut Variabel

Pada Bagian ini menjelaskan tentang atribut atribut yang digunakan dan juga label sebagai output pada prediksi

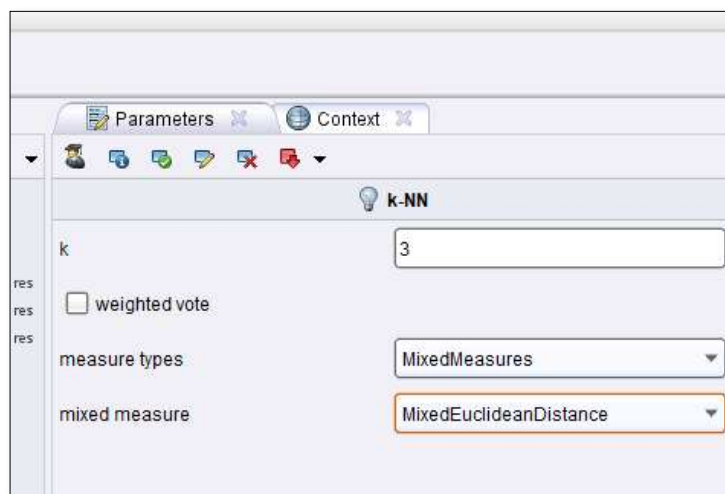
5.1.1.2. Pembagian Data Set



Gambar 5.4 Split Validation

Pada Pengaturan Dataset. Digunakan Model Split Validation dengan Persentase 90%, sehingga didapatkan Persentase Training sebesar 0,9 dan testing sebesar 0.1 atau perbangan data training 66 recod dan data testing 6 record data

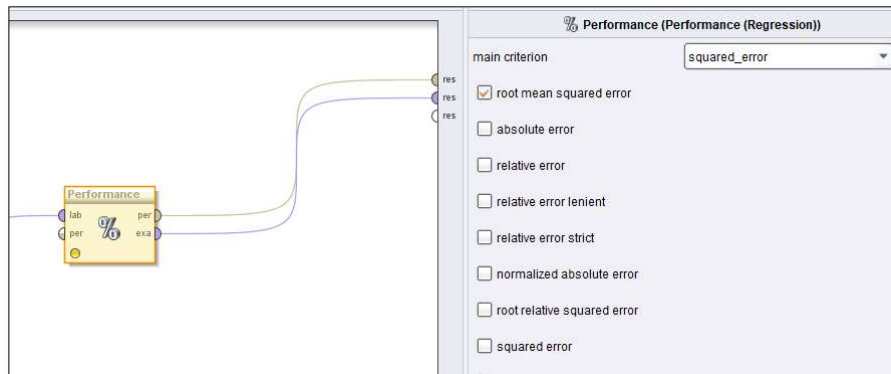
5.1.1.3. Properti K-NN



Gambar 5.5 Properti K-NN

Pada Penelitian Ini model k-NN yang digunakan adalah nilai K= 3 dan Perhitungan jarak dengan Euclidean Distance

5.1.1.4. Performace



Gambar 5.6 Performance

Untuk Mengetahui Kinerja dari Prediksi maka digunakan model Performace Untuk Kasus Regresion. Dengan main criterion *Root Mean Square Error* (RMSE)

5.1.1.5. Hasil Prediksi

ExampleSet (6 examples, 3 special attributes, 4 regular attributes)							
Row No.	TAHUN	BULAN	LUAS LAHAN	JUMLAH POKOK	CURAH H...	PRODUKSI	prediction(PRODUKSI)
1	2020	DESEMBER	3127.740	314.104	128.600	4656.060	4435.507
2	2021	JANUARI	3129.500	337.305	313.200	4417.820	3963.667
3	2021	FEBRUARI	3129.500	337.305	152.400	3745.230	3686.742
4	2021	MARET	3129.500	337.305	168.500	3946.400	3686.742
5	2021	APRIL	3129.500	337.305	142.200	3846.040	4408.570
6	2021	MEI	3129.500	337.305	218.600	2278.360	3140.266

Gambar 5.7 Hasil Prediksi

Gambar 5.7 Menampilkan Hasil Prediksi Untuk Data Testing Sebanyak 10 % sehingga didapatkan 6 Record data. Pada tabel tersebut terdapat perbandingan data sebenarnya (Aktual) pada kolom produksi dan Kolom Prediction (Produksi)

5.2. Pengujian Model *K-Nearest Neighbor (K-NN)*

Persoalan Regresi adalah proses identifikasi relasi dan pengaruhnya pada nilai-nilai objek. Regresi sendiri bertujuan untuk menemukan suatu fungsi yang memodelkan data dengan meminimalisir galat atau selisih antara nilai prediksi dengan nilai sebenarnya. Sedangkan untuk pengujian model algoritmanya adalah dengan menentukan nilai Root Mean Square Error (RMSE)

5.3 *RMSE (Root Mean Square Error)*

Root Mean Square Error (RMSE), adalah jumlah dari kesalahan kuadrat atau selisih antara nilai sebenarnya dengan nilai prediksi yang telah ditentukan. Rumus formula RMSE adalah sebagai berikut :

$$RMSE = \sqrt{\frac{\sum (Y' - Y)^2}{n}}$$

Y' = Nilai Prediksi

Y = Nilai Aktual

n = Jumlah Data

RMSE dapat berkisar dari 0 hingga ∞ . RMSE dapat berorientasi negatif dimana nilai yang lebih rendah menunjukkan nilai yang lebih baik. Jika nilai RMSE lebih kecil, hal ini berarti bahwa nilai yang diprediksi dekat dengan nilai yang diamati atau observasi, dan sebaliknya.

Tabel 5.1. RMSE Untuk Prediksi K=3

NO	Aktual (Y)	Prediksi (Y')	Error (Y-Y')	Error ²
66	4.020,79	4.262,18	-241,39	58.269,13
67	3.075,66	2.608,33	467,33	218.397,33
68	2.772,85	5414,878	-2.642,03	6.980.311,95
69	3.224,43	3.999,93	-775,50	601.400,25
70	3.535,62	3.882,07	-346,45	120.027,60
71	3.812,42	3.214,65	597,77	357.328,97

<i>NO</i>	<i>Aktual (Y)</i>	<i>Prediksi (Y')</i>	<i>Error (Y-Y')</i>	<i>Error ²</i>
72	3.639,93	3.443,36	196,57	38.639,76
Total				8.374.375
$\text{RMSE} = \sqrt{\text{total}/n}$ $= \sqrt{8.374.375/7}$ $= 1093,77$				

Tabel 5.2. RMSE Untuk Prediksi K=5

<i>NO</i>	<i>Aktual (Y)</i>	<i>Prediksi (Y')</i>	<i>Error (Y-Y')</i>	<i>Error ²</i>
66	4.020,79	4.168,01	-147,22	21.673,73
67	3.075,66	3.283,10	-207,44	43.031,35
68	2.772,85	3.554,18	-781,33	610.476,57
69	3.224,43	3.432,47	-208,04	43.280,64
70	3.535,62	3.882,07	-346,45	120.027,60
71	3.812,42	3.877,55	-65,13	4.241,92
72	3.639,93	3.674,81	-34,88	1.216,61
Total				843.948,43
$\text{RMSE} = \sqrt{\text{total}/n}$ $= \sqrt{8.374.375/7}$ $= 347,22$				

Tabel 5.3. RMSE Untuk Prediksi K=7

<i>NO</i>	<i>Aktual (Y)</i>	<i>Prediksi (Y')</i>	<i>Error (Y-Y')</i>	<i>Error ²</i>
66	4.020,79	4.324,15	-303,36	92.027,29
67	3.075,66	3.104,22	-28,56	815,67
68	2.772,85	3.452,12	-679,27	461.407,73
69	3.224,43	3.412,33	-187,90	35.306,41
70	3.535,62	3.896,21	-360,59	130.025,15
71	3.812,42	3.683,71	128,71	16.566,26

NO	Aktual (Y)	Prediksi (Y')	Error (Y-Y')	Error ²
72	3.639,93	3.514,45	125,48	15.745,23
Total				751.893,75
$\text{RMSE} = \sqrt{\text{Total}/n}$ $= \sqrt{751.893,75/7}$ $= 327,73$				

Nilai RMSE rendah menunjukkan bahwa variasi nilai yang dihasilkan oleh suatu model prakiraan mendekati variasi nilai obeservasinya. RMSE menghitung seberapa berbedanya seperangkat nilai. Semakin kecil nilai RMSE, semakin dekat nilai yang diprediksi dan diamati



Gambar 5.8 Hasil Pengujian dengan RMSE

RMSE merupakan kriteria penting dalam memilih model berperforma terbaik di antara model peramalan berbeda. Untuk melakukannya, cukup bandingkan nilai RMSE di semua model dan pilih satu dengan nilai RMSE terendah. Berdasarkan pengujian yang dilakukan model RMSE terbaik adalah menggunakan K=7 dengan RMSE 327,73.

BAB VI

PENUTUP

6.1 Kesimpulan

1. Berdasarkan hasil penelitian yang telah dilakukan dan pembahasan yang telah diuraikan pada sebelumnya maka didapatkan suatu kesimpulan bahwa Model telah diuji coba menggunakan aplikasi Rapid Miner untuk prediksi jumlah produksi buah kelapa sawit. Dimana dataset yang digunakan dalam bentuk excel. Pada Pengaturan Dataset. Digunakan Model Split Validation dengan Persentase 90%, sehingga didapatkan Persentase Training sebesar 0,9 dan testing sebesar 0.1 atau perbandingan data training 66 record dan data testing 6 record data. Hasil Prediksi Untuk Data Testing Sebanyak 10 % sehingga didapatkan 6 Record data. Data yang digunakan dalam penelitian ini adalah data histori dari Tahun 2016-2021 dengan menggunakan metode K-NN dengan hasil prediksi terdapat perbandingan data sebenarnya (Aktual) pada kolom produksi dan Kolom Prediction (Produksi). Ditetapan berdasarkan hasil produksi setiap bulannya dimana jika hasil produksi melebihi 5.000kg maka “Meningkat” dan jika kurang dari 5.000kg maka “Menurun”.
2. Model KNN Terbaik berdasarkan pengujian RMSE yang dilakukan terhadap uji Coba K=3, 5 dan 7, nilai RMSE terendah didapatkan pada pengujian K=7 dengan nilai 327,73

6.2 Saran

Dari Hasil Ekperimen yang telah dilakukan dan didapatkan model K-NN Terbaik maka peneliti menyarankan untuk membangun sistem prediksi dalam bentuk aplikasi dekstop maupun berbasis website.sehingga dapat diterapkan pada PT. HARDAYA INTI PLANTATIONS.

DAFTAR PUSTAKA

- [1] Limanseto H. Industri Kelapa Sawit Indonesia: Menjaga Keseimbangan Aspek Sosial, Ekonomi, dan Lingkungan. WwWekonGoId 2021.
- [2] BPS Kab. Buol. Luas Areal, Luas Panen, dan Produksi Tanaman Perkebunan menurut Kecamatan dan Jenis Tanaman Tahun 2013. BuolkabBpsGoId/ 2015:1–1.
- [3] Saadah S, Zahra F, Haifa H. Support Vector Regression (SVR) Dalam Memprediksi Harga Minyak Kelapa Sawit di Indonesia dan Nilai Tukar Mata Uang EUR/USD (Support Vector Machine (SVM) To Predict Crude Oil Palm in Indonesia and Exchange Rate of EUR/USD). n.d.
- [4] Agasta E, Cholissodin I, Ratnawati DE. Prediksi Jumlah Produksi Kelapa Sawit Dengan Menggunakan Metode Extreme Learning Machine (ELM) (Studi kasus: PT. Sandabi Indah Lestari Kota Bengkulu). vol. 2. 2018.
- [5] Kurniawan Sumba D. Sistem Cerdas Prediksi Hasil Tanaman Jagung Di Indonesia Menggunakan Support Vector Regression. Jurnal Nasional CosPhi 2018;2.
- [6] al Faraby S. Analisis Dan Implementasi Support Vector Machine Dengan String Kernel Dalam Melakukan Klasifikasi Berita Berbahasa Indonesia Analysis and Implementation Support Vector Machine With String Kernel for Classification indonesian news. Bandung: 2018.
- [7] Dewi AK, Furqon MT, Wihandika RC. Prediksi Laju Pertumbuhan Penduduk Menggunakan Metode Support Vector Regression (Studi Kasus: Kota Malang). vol. 4. 2020.
- [8] Lubis ER, Widanarko A. Buku Pintar Kelapa Sawit. vol. VIII. Cetak 1. Jakarta: PT. Agromedia Pustaka; 2011.
- [9] Fauzi Y, Widiastuti E Y, satyawibawa I, Paeru H R. Kelapa Sawit. 1st ed. Jakarta: Penebar Sawadaya Grup; 2012.
- [10] Arifin M. Implementasi Data Mining Pada Prediksi Pemesanan Menggunakan Algoritma Apriori (Studi Kasus : Kimia Farma). Jurnal Pelita Informatika 2020;8.

- [11] Kafil M. Penerapan Metode K-Nearest Neighbors Untuk Prediksi Penjualan Berbasis Web Pada Boutiq Dealove Bondowoso. vol. 3. 2019.
- [12] Sudirman I. Analisis Proses Pembentukan Kebijakan Publik: Studi Kasus Program Beasiswa Karawang Cerdas Tahun 2020. vol. 2. 2021.
- [13] HAFANDI A. Implementasi Data Mining Untuk Prediksi Daerah Rawan Penyakit Demam Berdarah Menggunakan Algoritma C4.5 (Studi Kasus : Dinas Kesehatan Kabupaten Tangerang). 2018.
- [14] Mahmud R, Hartanto A. Penerapan Data Mining Rekomendasi Laptop Menggunakan Algoritma Apriori. JUISI 2020;06.
- [15] Ni Lih Wiwik. DATA MINING DAN PENERAPAN ALGORITMA. Cetakan 1. Yayasan Kita Menulis; 2021.
- [16] Sukisno SKom, MK. PEMANFAATAN VECTOR SPACE MODEL. In: Hidayat T, editor. thesis. Cetakan Pertama, Jawa Tengah: CV. Pena Persada; 2021.
- [17] Teknomo K (2006). What is K-Nearest Neighbor Algoritma ? Dipetik 06 22, 2016 dari <http://people.nevolude.com/kardi/tutorial/KNN/What-is-K-Nearest-Neighbor-Algorithm.html>.
- [18] T. Chai and R. R. Draxler, 2014, “*Root Mean Square Error (RMSE) or Mean Absolute Error (MAE)? -Arguments Against Avoiding RMSE In The Literature*”, Vol 7.

Lampiran 1. Hasil Prediksi Menggunakan Raapid Miner

Hasil Prediksi

ExampleSet (6 examples, 3 special attributes, 4 regular attributes)							
Row No.	TAHUN	BULAN	LUAS LAHAN	JUMLAH POKOK	CURAH H...	PRODUKSI	prediction(PRODUKSI)
1	2020	DESEMBER	3127.740	314.104	128.600	4656.060	4435.507
2	2021	JANUARI	3129.500	337.305	313.200	4417.820	3963.667
3	2021	FEBRUARI	3129.500	337.305	152.400	3745.230	3686.742
4	2021	MARET	3129.500	337.305	168.500	3946.400	3686.742
5	2021	APRIL	3129.500	337.305	142.200	3846.040	4408.570
6	2021	MEI	3129.500	337.305	218.600	2278.360	3140.266

Gambar 5.7 Hasil Prediksi

Gambar 5.7 Menampilkan Hasil Prediksi Untuk Data Testing Sebanyak 10 % sehingga didapat 6 Record data. Pada tabel tersebut terdapat perbandingan data sebenarnya (Aktual) pada kolom produksi dan Kolom Prediction (Produksi)

Lampiran 2. Surat Keterangan Penelitian



PT. HARDAYA INTI PLANTATIONS

Head Office : Jl. Cikini Raya No. 78 Jakarta Pusat 10330 Telp. : 021 - 3149088, 314089 Fax. : 021 - 3150447
Branch Office : Jl. Dewi Sartika II, Lorong Jembolan No. 32, Palu - Sulawesi Tengah Telp & Fax : 0451 - 487659
Plantations : Desa Winangun, Kab. Buol - Sulawesi Tengah, E-mail : hip@cbn.net.id Website : www.hardaya.co.id

Kebun Modo, 7 September 2021

Nomor : **S23 /HIP/IP/IX/2021**
Lampiran : -
Perihal : **Izin Penelitian**

Kepada Yth.

Ketua Lembaga Penelitian Universitas
Ihsan Gorontalo

Tempat

Dengan hormat,

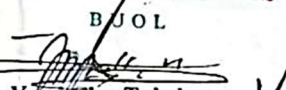
Menindaklanjuti surat Ketua Lembaga Penelitian Universitas Ihsan Gorontalo Nomor 1612/PIP/LEMLIT-UNISAN/GTO/VIII/2021 Tanggal, 07 Agustus 2021 yang kami terima tanggal, 31 Agustus 2021, dengan ini disampaikan bahwa :

1. Perusahaan memberi izin bagi mahasiswa atas nama **Sinta Lestari Aspa** akan melaksanakan kegiatan Penelitian selama 2 (dua) minggu, terhitung tanggal, 8 September s/d 18 September 2021, dengan judul penelitian : **Prediksi Jumlah Produksi Buah Kelapa Sawit Menggunakan Metode VCR.**
2. Selama melaksanakan Penelitian diwajibkan menjalankan protokol kesehatan.
3. Mahasiswa tersebut harus menyediakan dokumen, yaitu:
 - a. Surat Keterangan Bebas Covid-19 (Rapid Test).
 - b. Foto copy KTP 1 lembar.
 - c. Surat Pernyataan Bersedia Menjaga Kerahasiaan Data (tanda tangan di atas meterai 10.000).
 - d. Dokumen tersebut dimasukkan dalam map plastik snalheacter berwarna biru.

Demikian disampaikan terima kasih.

PT. HARDAYA INTI PLANTATIONS
PT. HARDAYA INTI PLANTATIONS

BUOL


Yussy Glen Tubalawony
Manager HRGA

Tembusan : Dengan hormat disampaikan kepada :

1. Mill Manager PMKS di Tempat
2. Ka.TU PMKS di Tempat.
3. Arsip

Lampiran 3. Surat Rekomendasi Bebas Pustaka



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

SURAT KETERANGAN BEBAS PUSTAKA

No : 023/Perpustakaan-Fikom/V/2022

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

Nama Anggota : Sinta Lestari Aspa
No. Induk : T3115202
No. Anggota : M202227

Terhitung mulai hari, tanggal : Sabtu, 28 Mei 2022, dinyatakan telah bebas pinjam buku dan koleksi perpustakaan lainnya.

Demikian keterangan ini di buat untuk di pergunakan sebagaimana mestinya.




Gorontalo, 28 Mei 2022

**Mengetahui,
Kepala Perpustakaan**

Apriyanto Alhamad, M.Kom
NIDN : 0924048601

Lampiran 4. Hasil Uji Turnitin

**Similarity Report ID:** oid:25211:18496430

PAPER NAME	AUTHOR
SKRIPSI_T3115202_SINTA LESTARI ASP A.docx	T3115202-SINTA LESTARI ASPA Shintap andu@gmail.com

WORD COUNT	CHARACTER COUNT
9520 Words	58798 Characters
PAGE COUNT	FILE SIZE
54 Pages	350.8KB
SUBMISSION DATE	REPORT DATE
Jun 9, 2022 9:54 PM GMT+8	Jun 9, 2022 9:57 PM GMT+8

● **24% Overall Similarity**
The combined total of all matches, including overlapping sources, for each database.

- 24% Internet database
- 5% Publications database
- Crossref database
- Crossref Posted Content database
- 0% Submitted Works database

● **Excluded from Similarity Report**

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

Lampiran 5. Riwayat Hidup

DAFTAR RIWAYAT HIDUP

Nama : Sinta lestari aspa
NIM : T3115202
Tempat Lahir : Kokobuka
Tanggal Lahir : 03 Juli 1997
Pekerjaan : Mahasiswa
Agama : Islam
Email : sintalestari030797@gmail.com



RIWAYAT PENDIDIKAN:

1. Tahun 2009, Menyelesaikan Pendidikan di Sekolah Dasar Negeri 7 Tiloan Ds.Kokobuka Kec.Tiloan Kab.Buol Provinsi Sulawesi Tengah.
2. Tahun 2012, Menyelesaikan Pendidikan di Sekolah MTS AL-Muhajirin Ds.Kokobuka Kec.Tiloan Kab.Buol Provinsi Sulawesi Tengah
3. Tahun 2015, Menyelesaikan Pendidikan di Sekolah Menengah Kejuruan Negeri 1 Biau Kec.Biau Kab.Buol Provinsi Sulawesi Tengah
4. Tahun 2015, Telah diterima Menjadi Mahasiswa di Perguruan Tinggi Swasta Universitas Ichsan Gorontalo.