

**PENERAPAN DATA *MINING* UNTUK PREDIKSI
PENJUALAN *SMARTPHONE* PALING LARIS
MENGUNAKAN METODE
K-NEAREST NEIGHBOR
(STUDI KASUS : PUSAT PONSEL & LAPTOP)**

TUGAS AKHIR

Disusun Untuk Memenuhi Syarat Kelulusan Program Strata I pada
Sekolah Tinggi Manajemen Informatika dan Komputer
(STMIK) Palangkaraya



OLEH

BRAEN DWIATMAJAYA GARANG
C1755201090
PROGRAM STUDI TEKNIK INFORMATIKA

**SEKOLAH TINGGI MANAJEMEN INFORMATIKA DAN KOMPUTER
(STMIK) PALANGKARAYA
2022**

**PENERAPAN DATA *MINING* UNTUK PREDIKSI
PENJUALAN *SMARTPHONE* PALING LARIS
MENGUNAKAN METODE
K-NEAREST NEIGHBOR
(STUDI KASUS : PUSAT PONSEL & LAPTOP)**

TUGAS AKHIR

Disusun Untuk Memenuhi Syarat Kelulusan Program Strata I pada
Sekolah Tinggi Manajemen Informatika dan Komputer
(STMIK) Palangkaraya

OLEH

BRAEN DWIATMAJAYA GARANG
C1755201090
PROGRAM STUDI TEKNIK INFORMATIKA

**SEKOLAH TINGGI MANAJEMEN INFORMATIKA DAN KOMPUTER
(STMIK) PALANGKARAYA
2022**

LEMBAR PERNYATAAN

Yang bertanda tangan di bawah ini:

Nama Mahasiswa : **BRAEN DWIATMAJAYA GARANG**
NIM : **C1755201090**

menyatakan bahwa Tugas Akhir dengan judul :

**PENERAPAN DATA *MINING* UNTUK PREDIKSI PENJUALAN
SMARTPHONE PALING LARIS MENGGUNAKAN METODE
K-NEAREST NEIGHBOR
(STUDI KASUS : PUSAT PONSEL & LAPTOP)**

adalah hasil karya saya dan bukan merupakan duplikasi sebagian atau seluruhnya dari karya orang lain, kecuali bagian yang sumber informasi dicantumkan.

Pernyataan ini dibuat dengan sebenar-benarnya secara sadar dan bertanggungjawab dan saya bersedia menerima sanksi pembatalan Tugas Akhir apabila terbukti melakukan duplikasi terhadap Tugas Akhir atau karya ilmiah lain yang sudah ada.

Palangka Raya, 29 Juni 2022

Yang Membuat Pernyataan,



BRAEN DWIATMAJAYA GARANG

PERSETUJUAN

PENERAPAN DATA *MINING* UNTUK PREDIKSI PENJUALAN *SMARTPHONE* PALING LARIS MENGUNAKAN METODE *K-NEAREST NEIGHBOR* (STUDI KASUS : PUSAT PONSEL & LAPTOP)

Tugas Akhir ini Telah Disetujui Untuk Diujikan pada
Tanggal 22 Juni 2022

Pembimbing I,



Veny Cahya Hardita, M.Kom.
NIK. 199504302020002

Pembimbing II,



Elok Faiqotul Himmah, S.Si., M.Sc.
NIK. 198503092009003

Mengetahui

Ketua STMIK Palangkaraya,



Suparno, M.Kom.

NIK. 196901041995105

PENGESAHAN

PENERAPAN DATA *MINING* UNTUK PREDIKSI PENJUALAN *SMARTPHONE* PALING LARIS MENGUNAKAN METODE *K-NEAREST NEIGHBOR* (STUDI KASUS : PUSAT PONSEL & LAPTOP)

Tugas Akhir ini Telah Diujikan, Dinilai, dan Disahkan
Oleh Tim Penguji pada Tanggal 25 Juni 2022

Tim Penguji Tugas Akhir :

1. Hotmian Sitohang, M.Kom.
Ketua



.....

2. Ir. Hj. Siti Maryamah, M.M.
Sekretaris



.....

3. Lili Rusdiana, M.Kom.
Anggota



.....

4. Veny Cahya Hardita, M.Kom.
Anggota



.....

5. Elok Faiqotul Himmah, S.Si., M.Sc.
Anggota



.....

MOTTO DAN PERSEMBAHAN

Saya percaya dengan terus berusaha dan kerja keras bisa meningkatkan kemampuan dan kualitas hasil pekerjaan.

Tugas Akhir ini saya persembahkan untuk :

- Keluarga saya yang tercinta, yang selalu mendukung, mendidik saya, dan membesarkan saya hingga sekarang.
- Teman-teman Teknik Informatika A dan B Angkatan 2017, yang sudah membantu saya.
- Dosen-dosen STMIK Palangkaraya, yang telah memberikan ilmu serta membimbing saya dari awal perkuliahan sampai lulus dan mendapatkan gelar sarjana komputer (S.Kom).

INTISARI

Braen Dwiatmajaya Garang, C1755201090, 2022. *Penerapan Data Mining Untuk Prediksi Penjualan Smartphone Paling Laris Menggunakan Metode K-Nearest Neighbor*, Pembimbing I Veny Cahya Hardita, M.Kom., Pembimbing II Elok Faiqotul Himmah, S.Si., M.Sc.

Penjualan merupakan sesuatu yang sangat penting bagi perusahaan yaitu memperoleh penghasilan dan laba untuk perusahaan agar tetap bisa berjalan. PPL (Pusat Ponsel & laptop) merupakan toko komputer yang menjual laptop, printer & *smartphone* baru di kota Palangka Raya. Untuk mempermudah toko dalam penyediaan stok barang maka diperlukan sebuah prediksi berdasarkan data-data penjualan terdahulu menggunakan metode *K-Nearest Neighbor*. Penelitian ini bertujuan untuk membuat prediksi penjualan *smartphone* paling laris di PPL dengan menggunakan metode *K-Nearest Neighbor*.

Jenis penelitian menggunakan penelitian kuantitatif dengan pendekatan deskriptif. Data yang digunakan dalam penelitian ini yaitu data penjualan *smartphone* selama 6 bulan dari Oktober 2021 sampai Maret 2022 dengan total penjualan sebanyak 1519 unit *smartphone* dengan teknik pengumpulan data seperti observasi, dokumentasi dan studi kepustakaan yang terkait dengan topik penelitian. Selanjutnya, data akan dianalisis dengan mengikuti proses tahapan dalam *Knowledge Discovery in Database (KDD)*, untuk menentukan informasi yang berguna serta pola-pola yang ada di dalam data.

Untuk memprediksi penjualan yang akan datang dengan *Euclidean Distance*, nilai $k = 3$ menggunakan aplikasi *Rapidminer* didapatkan hasil prediksinya untuk setiap *brand smartphone* dengan total penjualan oppo sebanyak 130 unit, vivo sebanyak 210 unit, samsung sebanyak 53 unit, xiaomi sebanyak 23 unit, dan realme sebanyak 44 unit *smartphone*. Hasil evaluasi algoritma yang diperoleh dari nilai *RMSE* pada setiap *brand smartphone* seperti oppo dengan nilai 6,643, vivo dengan nilai 11,468, samsung dengan nilai 3,538, xiaomi dengan nilai 4,23, dan realme dengan nilai 3,325. Dapat disimpulkan bahwa *smartphone* yang paling laris adalah vivo dan *brand smartphone* realme memiliki *RMSE* terkecil, yang berarti memiliki hasil prediksi yang mendekati dengan nilai yang sebenarnya.

Kata kunci : *Data Mining, K-Nearest Neighbor, Prediksi penjualan, RapidMiner.*

ABSTRACT

Braen Dwiatmajaya Garang, C1755201090, 2022. *Penerapan Data Mining Untuk Prediksi Penjualan Smartphone Paling Laris Menggunakan Metode K-Nearest Neighbor*, Pembimbing I Veny Cahya Hardita, M.Kom., Pembimbing II Elok Faiqotul Himmah, S.Si., M.Sc.

Sales is something that is very important for the company, namely earning income and profits for the company to keep it running. PPL (Center for Cellphones & Laptops) is a computer shop that sells new laptops, printers & smartphones in the city of Palangka Raya. To facilitate the store in providing stock, a prediction is needed based on previous sales data using the K-Nearest Neighbor method. This study aims to predict the best-selling smartphone sales in PPL using the K-Nearest Neighbor method.

This type of research uses quantitative research with a descriptive approach. The data used in this study is smartphone sales data for 6 months from October 2021 to March 2022 with a total sale of 1519 smartphone units with data collection techniques such as observation, documentation and literature studies related to the research topic. Furthermore, the data will be analyzed by following the stages in the Knowledge Discovery in Database (KDD) process, to determine useful information and patterns in the data.

To predict future sales with Euclidean Distance, the value of $k = 3$ using the Rapidminer application, the prediction results are obtained for each smartphone brand with a total sales of 130 units of Oppo, 210 units of Vivo, 53 Samsung units, 23 Xiaomi units, and Realme as many as 44 units of smartphones. The algorithm evaluation results obtained from the RMSE value on each smartphone brand such as Oppo with a value of 6.643, Vivo with a value of 11.468, Samsung with a value of 3.538, Xiaomi with a value of 4.23, and Realme with a value of 3.325. It can be concluded that the best-selling smartphone is vivo and the smartphone brand realme has the smallest RMSE, which means it has prediction results that are close to the actual value.

Keywords: Data Mining, K-Nearest Neighbor, Sales prediction, RapidMiner.

KATA PENGANTAR

Dengan mengucapkan syukur kepada Tuhan Yang Maha Esa, karena atas berkat dan rahmat-Nya penulis dapat menyelesaikan Tugas Akhir ini dengan judul “PENERAPAN DATA *MINING* UNTUK PREDIKSI PENJUALAN *SMARTPHONE* PALING LARIS MENGGUNAKAN METODE *K-NEAREST NEIGHBOR* (STUDI KASUS : PUSAT PONSEL & LAPTOP)”.

Penulis juga menyampaikan terima kasih yang sebesar-besarnya kepada semua pihak yang telah membantu dalam penulisan Tugas Akhir ini, yaitu kepada:

1. Bapak Syahidin, selaku kepala penjualan toko PPL yang telah mengizinkan penelitian ini.
2. Ibu Veny Cahya Hardita, M.Kom., selaku dosen Pembimbing I yang banyak memberikan saran dan pengetahuan dalam penyelesaian Tugas Akhir ini.
3. Ibu Elok Faiqotul Himmah, S.Si., M.Sc., selaku dosen Pembimbing II yang sudah banyak memberikan saran dan masukan dalam penulisan Tugas Akhir ini.

Penulis juga berharap penelitian yang sudah dilakukan ini dapat digunakan sebagai referensi agar penelitian selanjutnya menjadi lebih baik.

Palangka Raya, Juni 2022

Penulis

DAFTAR ISI

LEMBAR PERNYATAAN	ii
PERSETUJUAN	iii
PENGESAHAN	iv
MOTTO DAN PERSEMBAHAN	v
INTISARI	vi
ABSTRACT	vii
KATA PENGANTAR	viii
DAFTAR ISI	ix
DAFTAR TABEL	x
DAFTAR GAMBAR	xi
DAFTAR LAMPIRAN	xii
BAB I PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Perumusan Masalah	2
1.3 Batasan Masalah	3
1.4 Tujuan dan Manfaat	3
1.5 Sistematika Penulisan	4
BAB II LANDASAN TEORI	6
2.1 Tinjauan Pustaka	6
2.2 Kajian Teori	9
BAB III METODE PENELITIAN	22
3.1 Lokasi Penelitian.....	22
3.2 Perencanaan Alat dan bahan	22
3.3 Jenis Penelitian	23
3.4 Teknik Pengumpulan Data.....	23
3.5 Analisis	24
BAB IV HASIL DAN PEMBAHASAN	39
4.1 Pengujian Algoritma	39
4.2 Evaluasi Algoritma	46
BAB V KESIMPULAN DAN SARAN	48
5.1 Kesimpulan	48
5.2 Saran	49
DAFTAR PUSTAKA	
LAMPIRAN	

DAFTAR TABEL

Tabel 1. <i>Hardware</i> yang digunakan.....	22
Tabel 2. <i>Software</i> yang digunakan.....	23
Tabel 3. Data Penjualan <i>Smartphone</i> Oktober 2021.....	25
Tabel 4. Data Penjualan <i>Smartphone</i> November 2021.....	26
Tabel 5. Data Penjualan <i>Smartphone</i> Desember 2021.....	27
Tabel 6. Data Penjualan <i>Smartphone</i> Januari 2022.....	27
Tabel 7. Data Penjualan <i>Smartphone</i> Februari 2022.....	28
Tabel 8. Data Penjualan <i>Smartphone</i> Maret 2022.....	29
Tabel 9. Sampel Data Penjualan <i>Smartphone</i> Oppo.....	30
Tabel 10. Sampel Data Penjualan <i>Smartphone</i> Vivo.....	30
Tabel 11. Sampel Data Penjualan <i>Smartphone</i> Samsung.....	30
Tabel 12. Sampel Data Penjualan <i>Smartphone</i> Xiaomi.....	31
Tabel 13. Sampel Data Penjualan <i>Smartphone</i> Realme.....	31
Tabel 14. Data <i>Training</i> Penjualan <i>Smartphone</i> Oppo.....	31
Tabel 15. Data <i>Testing</i> Penjualan <i>Smartphone</i> Oppo.....	32
Tabel 16. Data <i>Training</i> Penjualan <i>Smartphone</i> Vivo.....	33
Tabel 17. Data <i>Testing</i> Penjualan <i>Smartphone</i> Vivo.....	33
Tabel 18. Data <i>Training</i> Penjualan <i>Smartphone</i> Samsung.....	33
Tabel 19. Data <i>Testing</i> Penjualan <i>Smartphone</i> Samsung.....	34
Tabel 20. Data <i>Training</i> Penjualan <i>Smartphone</i> Xiaomi.....	34
Tabel 21. Data <i>Testing</i> Penjualan <i>Smartphone</i> Xiaomi.....	34
Tabel 22. Data <i>Training</i> Penjualan <i>Smartphone</i> Realme.....	35
Tabel 23. Data <i>Testing</i> Penjualan <i>Smartphone</i> Realme.....	35
Tabel 24. Pengurutan hasil <i>Euclidean Distance</i>	37
Tabel 25. Nilai <i>RMSE</i> Untuk Semua <i>Brand Smartphone</i>	46

DAFTAR GAMBAR

Gambar 1. <i>Knowledge Discovery in Database</i>	15
Gambar 2. Rumus Perhitungan <i>Euclidean Distance</i>	18
Gambar 3. Rumus Perhitungan <i>RMSE</i>	38
Gambar 4. <i>Import Data</i>	39
Gambar 5. Data <i>Training</i> dan Data <i>Testing</i> di masukkan kedalam <i>Process</i>	40
Gambar 6. Operator <i>K-NN</i>	41
Gambar 7. Operator <i>Apply Model</i> dan <i>Performance</i>	41
Gambar 8. Susunan Operator	42
Gambar 9. Hasil Prediksi Penjualan <i>Smartphone</i> Oppo	43
Gambar 10. Hasil Prediksi Penjualan <i>Smartphone</i> Vivo	44
Gambar 11. Hasil Prediksi Penjualan <i>Smartphone</i> Samsung.....	44
Gambar 12. Hasil Prediksi Penjualan <i>Smartphone</i> Xiaomi	45
Gambar 13. Hasil Prediksi Penjualan <i>Smartphone</i> Realme	45

DAFTAR LAMPIRAN

- Lampiran 1. Surat tugas pembimbing Tugas Akhir.
- Lampiran 2. Lembar konsultasi bimbingan Tugas Akhir
- Lampiran 3. Surat izin penelitian
- Lampiran 4. Surat keterangan telah melakukan penelitian
- Lampiran 5. Surat tugas penguji sidang
- Lampiran 6. Berita acara penilaian sidang TA

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Pada masa sekarang ini perkembangan teknologi dan komunikasi dari waktu ke waktu dirasakan semakin meningkat pesat, yang mendorong penggunaan dan pemanfaatan perkembangan teknologi di berbagai bidang dan aspek kehidupan, contohnya seperti penggunaan *smartphone*. Masyarakat jaman sekarang bahkan kalangan menengah kebawah bisa memiliki *smartphone*. Perkembangan teknologi ini memicu persaingan oleh berbagai *brand smartphone* yang ada dalam melakukan penjualan, sehingga membuat tiap *brand smarphone* harus memiliki fitur dan performa yang lebih baik agar tetap bisa bersaing.

Penjualan merupakan sesuatu yang sangat penting bagi perusahaan yaitu memperoleh penghasilan dan laba untuk perusahaan agar tetap bisa berjalan. PPL (Pusat Ponsel & laptop) merupakan toko komputer yang menjual laptop, printer & *smartphone* baru, serta *service* laptop, printer, dan *smartphone* di kota Palangka Raya. PPL juga melayani kredit *smartphone* dan kredit laptop untuk keperluan anak sekolah, anak kuliah, dan kerja sehari-hari. Permasalahan yang umum biasanya seperti bagaimana penjualan barang di masa yang akan datang lebih laku. Untuk mempermudah toko dalam penyediaan stok barang maka diperlukan sebuah prediksi berdasarkan data-data penjualan terdahulu menggunakan data *mining*.

Data *mining* adalah bidang dari beberapa bidang keilmuan yang menyatukan teknik dari pembelajaran mesin, pengenalan pola, statistik, *database* dan visualisasi untuk penanganan permasalahan informasi dari penyimpanan *database* yang besar (Amalia, 2018). Data *mining* dapat digunakan untuk menggali informasi dari data yang besar sehingga didapatkan informasi yang dapat digunakan dalam peramalan penjualan. Peramalan penjualan merupakan suatu perhitungan untuk meramalkan keadaan di masa mendatang. Ada banyak metode peramalan penjualan, dan penulis memutuskan untuk menggunakan metode *K-Nearest Neighbor*. Metode ini merupakan metode klasifikasi yang bisa digunakan dalam data *mining* untuk melakukan prediksi dengan menghitung jarak terdekat dalam data. Prediksi tersebut dapat digunakan untuk melakukan perencanaan stok barang yang lebih banyak peminatnya sehingga perusahaan bisa meraih untung. Berdasarkan latar belakang diatas, penulis tertarik untuk mengambil sebuah judul “Penerapan Data *Mining* Untuk Prediksi Penjualan *Smartphone* Paling Laris Menggunakan Metode *K-Nearest Neighbor*”.

1.2 Perumusan Masalah

Berdasarkan uraian latar belakang diatas, maka perumusan masalahnya adalah bagaimana cara Memprediksi Penjualan *Smartphone* Paling Laris Menggunakan Metode *K-Nearest Neighbor*?

1.3 Batasan Masalah

Batasan-batasan masalah sebagai berikut:

- a. Prediksi jumlah penjualan *smartphone* berdasarkan data penjualan dari bulan Oktober 2021 sampai Maret 2022 yang ada di Pusat Ponsel & Laptop (PPL) Palangka Raya.
- b. Data yang digunakan yaitu data penjualan di PPL Palangka Raya selama 6 bulan dari bulan Oktober 2021 sampai Maret 2022 dengan total penjualan sebanyak 1519 unit *smartphone* dengan *brand* oppo, vivo samsung, xiaomi dan realme.
- c. Penelitian ini menggunakan metode *K-Nearest Neighbor* dan *tools* bantuan yaitu *RapidMiner*.

1.4 Tujuan dan Manfaat

- a. Tujuan

Tujuan penelitian ini adalah untuk membuat prediksi jumlah penjualan *smartphone* paling laris menggunakan metode *K-Nearest Neighbor* pada toko Pusat Ponsel & Laptop.

- b. Manfaat

Penelitian ini diharapkan dapat memberikan manfaat, yaitu:

- 1) Bagi penulis

Manfaat bagi penulis adalah menambah wawasan serta mengembangkan ilmu pengetahuan tentang penggunaan dan implementasi data *mining*.

2) Bagi STMIK Palangkaraya

Manfaat bagi pihak kampus adalah bisa digunakan sebagai referensi dalam bidang data *mining* dan pustaka pada perpustakaan STMIK Palangkaraya.

3) Bagi Perusahaan

Manfaat bagi perusahaan adalah dapat memprediksi jumlah penjualan *smartphone* yang paling laris agar bisa menyediakan stok barang yang lebih banyak untuk meraih keuntungan.

1.5 Sistematika Penulisan

Adapun sistematika penulisan dalam penulisan tugas akhir ini terdiri dari beberapa bab dan masing-masing bab membahas dan menguraikan pokok permasalahan yang berbeda, sebagai gambaran disini penulis menyertakan garis-garis besarnya yaitu:

BAB I PENDAHULUAN

Bab ini berisikan latar belakang masalah, rumusan masalah, batasan masalah, tujuan, manfaat, serta sistematika penulisan.

BAB II LANDASAN TEORI

Bab ini berisikan tentang tinjauan pustaka yang diambil dari penelitian yang relevan berserta susuna kajian teori, seperti *retail*, *smartphone*, data, teknik pengumpulan data, pengetahuan, data *mining*, *Knowledge Discovery in Database*, peramalan, metode *K-Nearest Neighbor*, dan *RapidMiner*.

BAB III METODE PENELITIAN

Bab ini berisikan tentang yang dilakukan peneliti dalam mengumpulkan informasi atau data yang dibutuhkan untuk menganalisis dan memprediksi penjualan *smartphone* berdasarkan data penjualan selama 6 bulan dari Oktober 2021 sampai Maret 2022.

BAB IV HASIL DAN PEMBAHASAN

Bab ini berisikan tentang hasil dari penelitian dan pembahasan dari hasil yang didapat seperti cara penggunaan aplikasi *rapidminer* untuk memprediksi penjualan *smartphone* menggunakan metode *K-Nearest Neighbor* berdasarkan *brand smartphone* seperti oppo, vivo, samsung, xiaomi, dan realme.

BAB V KESIMPULAN DAN SARAN

Bab ini berisikan tentang kesimpulan dan saran yang diberikan untuk pengembangan selanjutnya.

BAB II

LANDASAN TEORI

2.1 Tinjauan Pustaka

Dalam suatu penelitian diperlukan dukungan hasil-hasil penelitian yang telah ada sebelumnya yang berkaitan dengan penelitian serupa. Berikut hasil-hasil penelitian yang relevan dan pembahasan penelitian yang telah ada sebelumnya yang serupa dengan penelitian yang sedang dilakukan antara lain:

- a. Penelitian yang relevan dilakukan oleh Amalia (2018) berjudul “Penerapan Data *Mining* Untuk Prediksi Penjualan Produk Elektronik Terlaris Menggunakan Metode *K-Nearest Neighbor*” Untuk mengetahui penjualan produk elektronik terlaris digunakan teknik klasifikasi data *mining* dan algoritma *K-Nearest Neighbor*. Hasil dari penelitian ini adalah prediksi penjualan elektronik terlaris sebanyak 6 jenis produk dari 22 jenis produk yang terjual yaitu CTV, Lemari Es, DVD, Speaker, Mesin Cuci dan LCD. Berdasarkan nilai akurasi terhadap klasifikasi penjualan produk terlaris sebesar 92,51 %.
- b. Penelitian yang relevan dilakukan oleh WPR, et al. (2021) berjudul “Prediksi Penjualan Produk Unilever Menggunakan Metode *K-Nearest Neighbor*” Untuk mengetahui penjualan produk Unilever terlaris digunakan teknik klasifikasi data *mining* dan algoritma *K-Nearest Neighbor*. Hasil dari penelitian ini berdasarkan nilai akurasi tertinggi

terhadap klasifikasi penjualan produk sebesar 86,66% dan nilai akurasi terhadap klasifikasi penjualan produk sebesar 40%.

- c. Penelitian yang relevan dilakukan oleh Putri (2021) berjudul “Penerapan Data Mining Untuk Memprediksi Penjualan Buah dan Sayur Menggunakan Metode *K-Nearest Neighbor*” PT. Central Brastagi Utama merupakan supermarket yang menjual banyak produk konsumsi diantaranya buah dan sayur. Sayangnya selama ini belum ada sistem yang mengatur prediksi atau peramalan untuk penjualan buah dan sayur pada PT. Central Brastagi Utama. Sehingga sering terjadi penumpukan barang, barang rusak dan busuk, atau bahkan kekurangan stok barang yang mengakibatkan kerugian bagi perusahaan. Untuk mengatasi masalah tersebut diperlukan suatu peramalan atau *forecasting*. Maka digunakan teknik data *mining* dengan metode *K-Nearest Neighbor*. Diharapkan dengan menggunakan teknik tersebut dapat mengolah data 3 tahun terakhir menjadi suatu informasi yang dapat membantu pihak perusahaan dalam penyediaan stok barang. Hasil dari penelitian ini akan memprediksi penjualan sejumlah produk, dan menentukan kategori penjualan produk tersebut laris, sedang atau sedikit.
- d. Penelitian yang relevan dilakukan oleh Yolanda dan Fahmi (2021) berjudul “Penerapan Data Mining Untuk Prediksi Penjualan Produk Roti Terlaris Pada PT.Nippon Indosari Corpindo Tbk Menggunakan Metode *K-Nearest Neighbor*” agar mempermudah pihak perusahaan dalam memproduksi roti mana yang paling banyak diproduksi. Karena saat ini

sistem yang digunakan masih manual dengan itu data yang didapat kurang akurat dan efisien. Maka untuk mengetahui penjualan produk roti terlaris dibutuhkanlah sebuah data *mining* untuk memecahkan sebuah masalah dengan menggunakan metode *k-nearest neighbor*. Hasil penelitian ini adalah sebuah sistem yang membantu untuk memprediksi penjualan produk roti terlaris pada PT.Nippon Indosari Corpindo Tbk agar data yang didapat akurat dan efisien. Dengan menggunakan kriteria kuantitas produk dan kuantitas terjual.

- e. Penelitian yang relevan dilakukan oleh Dewi, et al. (2022) berjudul “Penerapan Data *Mining* Untuk Prediksi Penjualan Produk Terlaris Menggunakan Metode *K-Nearest Neighbor*” agar mempermudah pihak usaha dagang dalam perencanaan penyedia stok. Dikrenakan sistem yang sedang berjalan saat ini masih manual, untuk itu data yang didapat kurang akurat dan efisien. Maka untuk mengatasi hal ini, diperlukan sebuah sistem prediksi penjualan produk terlaris dengan teknik data *mining* yang menggunakan metode *k-nearest neighbor*. Penelitian ini menghasilkan sebuah sistem algoritma *k-nearest neighbor* pada teknik data *mining* yang membantu untuk memprediksi penjualan produk telaris pada UD Andar.

2.2 Kajian Teori

a. *Retail*

Retail adalah semua kegiatan yang terlibat dalam penjualan barang atau jasa secara langsung kepada konsumen akhir untuk penggunaan pribadi non bisnis (Kotler & Keller, 2016). *Retail* adalah kegiatan bisnis yang terlibat dalam menjual barang dan jasa kepada konsumen untuk pribadi, keluarga, atau keperluan rumah tangga mereka (Berman & Evans, 2012).

b. *Smartphone*

Telepon cerdas (*smartphone*) adalah telepon genggam yang mempunyai kemampuan tingkat tinggi, dengan fungsi yang menyerupai komputer. Bagi beberapa orang, telepon pintar merupakan telepon yang bekerja menggunakan seluruh perangkat lunak sistem operasi yang menyediakan hubungan standar dan mendasar bagi pengembang aplikasi. Bagi yang lainnya, telepon cerdas hanyalah merupakan sebuah telepon yang menyajikan fitur canggih seperti surel (surat elektronik), internet dan kemampuan membaca buku elektronik (*e-book*) atau terdapat papan ketik (baik sebagaimana jadi maupun terhubung keluar) dan penyambung VGA (Kaswandari, 2017). Dengan kata lain, telepon cerdas merupakan komputer kecil yang mempunyai kemampuan sebuah telepon yang mempunyai daya guna bagi manusia.

c. Data

Data adalah suatu istilah majemuk yang berarti fakta atau bagian dari fakta yang mengandung arti yang dihubungkan dengan kenyataan, simbol-simbol, gambar-gambar, angka-angka, huruf-huruf, atau simbol-simbol yang menunjukkan suatu ide, objek, kondisi atau situasi, dan lain-lain (Longkutoy, 2012). Pembagian data menurut sumber pengambilannya, dapat dibedakan atas dua jenis, yaitu:

- 1) Data Primer: yaitu data yang diperoleh atau dikumpulkan oleh orang yang melakukan penelitian atau yang bersangkutan yang memerlukannya. Data primer disebut juga data asli atau data baru.
- 2) Data Sekunder: yaitu data yang diperoleh atau dikumpulkan dari sumber-sumber yang telah ada. Data itu biasanya diperoleh dari perpustakaan atau laporan-laporan/ dokumen peneliti yang terdahulu. Data sekunder disebut juga data tersedia.

Pembagian data menurut sifatnya, dibagi menjadi dua yaitu:

- 1) Data Kualitatif: yaitu data yang tidak berbentuk bilangan. Data kualitatif berbentuk pernyataan verbal, simbol atau gambar. Contoh: warna, jenis kelamin, status perkawinan, dll.
- 2) Data Kuantitatif: yaitu data yang berbentuk bilangan, atau data kualitatif yang diangkakan. Contoh: tinggi, umur, jumlah, skor hasil belajar, temperatur, dll.

d. Teknik Pengumpulan Data

Menurut Hardani (2020) teknik pengumpulan data merupakan langkah yang paling strategis dalam penelitian, karena tujuan utama dari penelitian adalah mendapatkan data. Tanpa mengetahui teknik pengumpulan data, maka peneliti tidak akan mendapatkan data yang memenuhi standar data yang ditetapkan. secara umum terdapat empat macam teknik pengumpulan data, yaitu:

- 1) Observasi: merupakan suatu metode pengumpulan data dengan cara pengamatan langsung ke lapangan pada objek yang akan diteliti. Langkah awal, peneliti akan melakukan observasi langsung ke lokasi penelitian di Pusat Ponsel & Laptop Palangka Raya. Kegiatan observasi meliputi kondisi tempat penelitian serta kegiatannya.
- 2) Wawancara: merupakan tanya jawab lisan antara dua orang atau lebih secara langsung atau percakapan dengan maksud tertentu. Percakapan itu dilakukan oleh dua pihak, yaitu pewawancara (*interviewer*) yang mengajukan pertanyaan dan yang diwawancarai (*interviewee*) yang memberikan jawaban atas pertanyaan itu.
- 3) Dokumentasi: merupakan suatu metode pengumpulan data dengan melihat atau menganalisis dokumen-dokumen yang dibuat oleh subjek sendiri atau oleh orang lain tentang subjek. Sebagai langkah awal, peneliti akan bertanya dan meminta izin kepada pihak Pusat Ponsel & Laptop Palangka Raya untuk meminta data yang diperlukan seperti data – data penjualan *smartphone* dalam bentuk file excel.

- 4) Studi Kepustakaan: merupakan suatu metode dalam mempelajari berbagai literatur yang relevan dari buku, artikel ilmiah, berita, maupun dari internet yang terkait dengan topik penelitian. Sebagai langkah awal, peneliti akan mempelajari tentang data *mining* dan *rapidminer* dari literatur yang ada dari buku, jurnal penelitian dan internet.

e. Pengetahuan

Menurut Yuliana (2017) Pengetahuan seseorang terhadap objek mempunyai intensitas yang berbeda-beda, dan menjelaskan bahwa ada enam tingkatan pengetahuan yaitu sebagai berikut:

- 1) Pengetahuan (*knowledge*) diartikan hanya sebagai *recall* (ingatan). Seseorang dituntut untuk mengetahui fakta tanpa dapat menggunakannya.
- 2) Pemahaman (*comprehension*) yaitu memahami suatu objek bukan sekedar tahu, tidak sekedar dapat menyebutkan, tetapi harus dapat menginterpretasikan secara benar tentang objek yang diketahui.
- 3) Penerapan (*application*) diartikan apabila orang yang telah memahami objek tersebut dapat menggunakan dan mengaplikasikan prinsip yang diketahui pada situasi yang lain.
- 4) Analisis (*analysis*) yaitu kemampuan seseorang untuk menjabarkan dan memisahkan, kemudian mencari hubungan antara komponen-komponen yang terdapat dalam suatu objek.

- 5) Sintesis (*synthesis*) yaitu suatu kemampuan untuk menyusun formulasi baru dari formulasi-formulasi yang telah ada. Sintesis menunjukkan suatu kemampuan seseorang untuk merangkum atau meletakkan dalam suatu hubungan yang logis dari komponen-komponen pengetahuan yang dimiliki.
- 6) Penilaian (*evaluation*) yaitu suatu kemampuan seseorang untuk melakukan penilaian terhadap suatu objek tertentu didasarkan pada suatu kriteria atau norma-norma yang berlaku di masyarakat.

f. Informasi

Informasi adalah hasil pengolahan data yang memberikan arti dan manfaat (Susanto, 2013). Informasi adalah data yang telah diklasifikasikan atau diolah atau diinterpretasikan untuk digunakan dalam proses pengambilan keputusan (Sutabri, 2012).

g. Data Mining

Data *mining* adalah proses untuk mendapatkan informasi yang berguna dari basis data yang besar dan perlu diekstraksi agar menjadi informasi baru dan dapat membantu dalam pengambilan keputusan (Suntoro, 2019). Data *mining* adalah proses menganalisa data dari yang berbeda dan menyimpulkannya menjadi informasi atau pengetahuan atau pola yang penting untuk meningkatkan keuntungan, memperkecil biaya pengeluaran, atau bahkan keduanya (Witten, 2016).

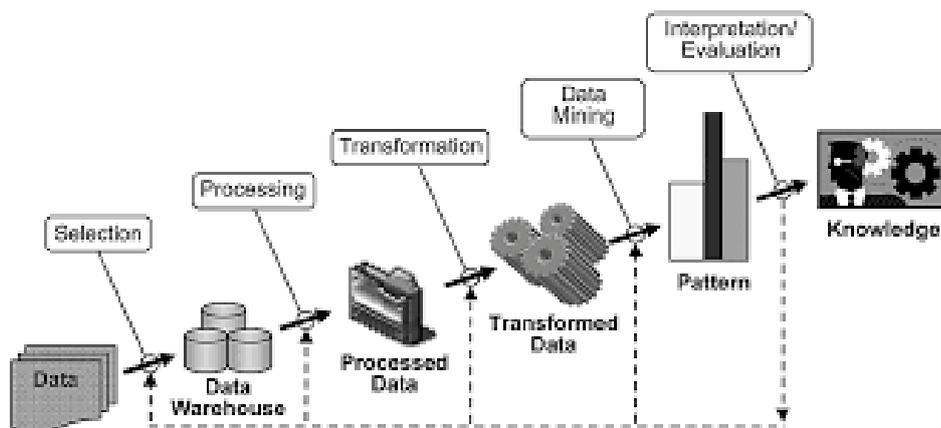
Menurut Rerung (2018) data *mining* dibagi menjadi beberapa kelompok berdasarkan tugas/ pekerjaan yang dapat dilakukan yaitu :

- 1) Deskripsi: terkadang peneliti dan analisis secara sederhana ingin mencoba mencari cara untuk menggambarkan pola dan kecenderungan yang terdapat dalam data. Deskripsi dari pola kecenderungan sering memberikan kemungkinan penjelasan untuk suatu pola atau kecenderungan.
- 2) Estimasi: hampir sama dengan klasifikasi, kecuali variabel target estimasi lebih ke arah numerik dari pada ke arah kategori. Model dibangun menggunakan baris data (*record*) lengkap yang menyediakan nilai dari variabel target sebagai nilai prediksi. Selanjutnya, pada peninjauan berikutnya estimasi nilai dari variabel target dibuat berdasarkan nilai variabel prediksi.
- 3) Prediksi: hampir sama dengan klasifikasi dan estimasi, kecuali bahwa dalam prediksi nilai dari hasil akan ada di masa mendatang. Beberapa metode dan teknik yang digunakan dalam klasifikasi dan estimasi dapat pula digunakan (untuk keadaan yang tepat) untuk prediksi.
- 4) Klasifikasi: terdapat target variabel kategori. Sebagai contoh, penggolongan pendapatan dapat dipisahkan dalam tiga kategori, yaitu pendapatan tinggi, pendapatan sedang, dan pendapatan rendah.
- 5) Pengklasteran: merupakan pengelompokan *record*, pengamatan, atau memperhatikan dan membentuk kelas objek-objek yang memiliki kemiripan. Klaster juga merupakan kumpulan *record* yang memiliki kemiripan satu dengan yang lainnya dan memiliki ketidakmiripan *record* dalam klaster yang lain. Berbeda dengan klasifikasi, pada

pengklasteran tidak ada variabel target. Pengklasteran tidak melakukan klasifikasi, mengestimasi, atau memprediksi nilai dari variabel target, akan tetapi, algoritma pengklasteran mencoba untuk melakukan pembagian terhadap keseluruhan data menjadi kelompok-kelompok yang memiliki kemiripan (homogen), yang mana kemiripan *record* dalam satu kelompok akan bernilai maksimal, sedangkan kemiripan dengan *record* dalam kelompok lain akan bernilai minimal.

- 6) Asosiasi: Tugas asosiasi dalam data *mining* yaitu untuk menemukan hubungan antara atribut yang muncul dalam satu waktu.

h. *Knowledge Discovery in Database*



Gambar 1. *Knowledge Discovery in Database*

Sumber: Amalia, 2018

Menurut Amalia (2018) *KDD* atau *Knowledge Discovery in Database* adalah sekumpulan proses yang digunakan untuk menemukan pengetahuan yang meliputi:

- 1) *Data Selection*: Pada proses ini dilakukan pemilihan himpunan data, menciptakan himpunan data target atau memfokuskan pada *subset*

variabel (sampel data) dimana penemuan (*discovery*) akan dilakukan. Hasil seleksi disimpan dalam suatu berkas yang terpisah dari basis data operasional.

- 2) *Preprocessing* dan *Cleaning Data*: dilakukan pembuangan data yang tidak konsisten dan *noise*, duplikasi data, memperbaiki kesalahan data, dan bisa diperkaya dengan data eksternal yang relevan.
- 3) *Transformation*: proses ini mentransformasikan atau menggabungkan data ke dalam yang lebih tepat untuk melakukan proses mining dengan cara melakukan peringkasan (agregasi).
- 4) *Data Mining*: yaitu proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik, metode atau algoritma tertentu sesuai dengan tujuan dari proses *KDD* secara keseluruhan.
- 5) *Interpretation* / Evaluasi: proses untuk menerjemahkan pola-pola yang dihasilkan dari data *mining*. Melakukan evaluasi (menguji) apakah pola atau informasi yang ditemukan bersesuaian atau bertentangan dengan fakta atau hipotesa sebelumnya. Pengetahuan yang diperoleh dari pola-pola yang terbentuk dipresentasikan dalam bentuk visualisasi.

a) *Estimation*:

Error: Root Mean Square Error (RMSE), MSE, MAPE, dll

b) *Prediction/Forecasting* (Prediksi/Peramalan):

Error: Root Mean Square Error (RMSE), MSE, MAPE, dll

c) *Classification:*

(1) *Confusion Matrix: Accuracy*

(2) *ROC Curve: Area Under Curve (AUC)*

d) *Clustering:*

(1) *Internal Evaluation: Davies–Bouldin index, Dunn index*

(2) *External Evaluation: Rand measure, F-measure, Jaccard index, Fowlkes–Mallows index, Confusion matrix*

e) *Association:*

(1) *Lift Charts: Lift Ratio*

(2) *Precision and Recall (F-measure)*

i. Peramalan

Peramalan (*forecasting*) adalah suatu proses untuk membuat informasi tentang situasi sosial masa depan atas dasar informasi yang terdapat pada situasi di masa lalu (Putri, 2021). Peramalan juga merupakan suatu proses perkiraan secara sistematis tentang sesuatu yang paling mungkin terjadi di masa depan berdasarkan informasi masa lalu dan informasi sekarang yang dimiliki, agar kesalahannya (selisih antara sesuatu yang terjadi dengan hasil perkiraan) dapat diperkecil. Prediksi merupakan hasil dari kegiatan memprediksi atau meramal atau memperkirakan nilai pada masa yang akan datang dengan menggunakan data masa lalu. Proses peramalan atau prediksi bisa berdasarkan metode ilmiah ataupun subjektif belaka.

j. Metode *K-Nearest Neighbor*

Algoritma *K-Nearest Neighbor* adalah pendekatan untuk mencari kasus dengan menghitung kedekatan antara kasus baru (*testing data*) dengan kasus lama (*training data*), yaitu berdasarkan pada pencocokan bobot dari sejumlah fitur yang ada (Amalia, 2018).

K-Nearest Neighbor (*K-NN*) adalah suatu metode yang menggunakan algoritma *supervised* dimana hasil dari *query instance* yang baru diklasifikasikan berdasarkan mayoritas dari *label class* pada *K-NN* (Yanosma, 2016). Pada algoritma *KNN* terdapat 5 (lima) cara untuk mencari tetangga terdekat yaitu:

- 1) Jarak *Euclidean*
- 2) Jarak *Manhattan*
- 3) Jarak *Cosine*
- 4) Jarak *Correlation*
- 5) Jarak *Hamming*

Pada penelitian ini akan menggunakan jarak *Euclidean*, maka rumus perhitungan jarak dengan *Euclidean* seperti persamaan 1:

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (1)$$

x_i dan y_i merupakan subjek yang akan dibandingkan sebanyak n . Dimana nilai x_i merupakan nilai yang ada pada data *training*, sedangkan nilai y_i merupakan nilai yang ada pada data *testing*, nilai n merupakan banyaknya data *testing*.

Parameter k merupakan nilai yang menentukan berapa banyak tetangga yang akan dipilih untuk algoritma KNN . Pilihan k yang tepat berdampak signifikan pada kinerja diagnostik algoritma $K-NN$. K besar mengurangi dampak varians yang disebabkan oleh kesalahan acak, tetapi berisiko mengabaikan pola kecil tapi penting. Kunci untuk memilih nilai k yang sesuai dengan mencapai keseimbangan antara *overfitting* dan *underfitting*.

Langkah-langkah untuk menghitung algoritma $K-NN$:

- 1) Menentukan nilai k (Jumlah tetangga terdekat). Umumnya nilai k ditentukan dengan nilai ganjil seperti 3, 5, 7, 9, dst.
- 2) Menghitung kuadrat jarak euclidean masing-masing objek terhadap *training* data yang diberikan.
- 3) Kemudian mengurutkan objek-objek tersebut ke dalam kelompok yang mempunyai jarak euclidean terkecil.
- 4) Mengumpulkan label target.
- 5) Dengan menggunakan kategori *Nearest Neighbor* yang paling mayoritas maka dapat diprediksikan nilai yang telah dihitung.

k. *RapidMiner*

RapidMiner adalah *platform* perangkat lunak data ilmu pengetahuan yang dikembangkan oleh sebuah perusahaan dengan nama yang sama, yang menyediakan lingkungan terpadu untuk pembelajaran mesin, pembelajaran mendalam, penambangan teks, dan analisis prediktif. Aplikasi ini digunakan untuk aplikasi bisnis dan komersial serta untuk

penelitian, pendidikan, pelatihan, pembuatan *prototype* dengan cepat, dan pengembangan aplikasi serta mendukung semua langkah proses pembelajaran mesin termasuk persiapan data, visualisasi hasil, validasi dan pengoptimalan (Rizqifaluthi & Yaqin, 2018).

RapidMiner adalah sebuah *software* pengolah data dengan menggunakan prinsip algoritma data *mining*. Cara kerja dari *rapidminer* ini dengan mengekstrak pola-pola data *set* yang berjumlah besar dengan cara mengkombinasikan metode statistika, kecerdasan buatan dan *database*. *Rapidminer* juga memudahkan pengguna untuk melakukan perhitungan data yang berjumlah besar dengan menggunakan operator yang berfungsi untuk memodifikasi data. *Rapidminer* menjadi salah satu perangkat lunak pilihan untuk mengekstraksi data dengan metode-metode data *mining* (Rahmat, et al., 2017).

RapidMiner menyediakan *GUI* (*Graphic User Interface*) untuk merancang sebuah *pipeline* analitis. *GUI* ini akan menghasilkan *file XML* (*Extensible Markup Language*) yang mendefinisikan proses analitis keinginan pengguna untuk diterapkan ke data. File ini kemudian dibaca oleh *RapidMiner* untuk menjalankan analisis secara otomatis.

RapidMiner memiliki beberapa sifat sebagai berikut:

- 1) Ditulis dengan bahasa pemrograman *java* sehingga dapat dijalankan di berbagai sistem operasi.
- 2) Proses penemuan pengetahuan dimodelkan sebagai operator trees.

- 3) Representasi *XML* internal untuk memastikan format standar pertukaran data.
- 4) Bahasa *scripting* memungkinkan untuk eksperimen skala besar dan otomatisasi eksperimen.
- 5) Konsep *multi-layer* untuk menjamin tampilan data yang efisien.
- 6) Memiliki *GUI*, *command line mode* dan *Java API* yang dapat dipanggil dari program lain.

Beberapa fitur dari *RapidMiner*, antara lain:

- 1) Banyaknya *algoritma* data mining, seperti *decision tree* dan *self-organization map*.
- 2) Bentuk grafis yang canggih, seperti tumbang tindih *diagram histogram*, *tree chart* dan *3D scatter plots*.
- 3) Banyaknya variasi *plugin*, seperti *text plugin* untuk melakukan analisis teks.
- 4) Menyediakan prosedur data mining dan *machine learning* termasuk: *ETL (extraction, transformation, loading)* data *preprocessing*, *visualisasi*, *modeling* dan *evalualisasi*.
- 5) Proses data mining tersusun atas operator-operator yang *nestable*, dideskripsikan dengan *XML*, dan dibuat dengan *GUI*.
- 6) Mengintegrasikan proyek data mining *Weka* dan *statistic R*.

BAB III

METODE PENELITIAN

3.1 Lokasi Penelitian

Lokasi Penelitian berada di toko Pusat Ponsel & Laptop (PPL) samping *traffic light*, Jl. Yos Sudarso Jl. Galaxy I No.6J, Palangka, Kec. Jekan Raya, Kota Palangka Raya, Kalimantan Tengah 73112.

3.2 Perencanaan Alat dan bahan

Dalam melakukan penelitian ini diperlukan suatu hardware dan *software* untuk menjalankan penelitian, berikut Tabel 1 dan Tabel 2 menunjukkan *hardware* dan *software* yang akan digunakan dalam penelitian ini.

Tabel 1. *Hardware* yang digunakan

No	Nama Perangkat	Spesifikasi
1	<i>Processor</i>	<i>Intel Core i5</i>
2	<i>RAM</i>	<i>4 GB</i>
3	<i>Harddisk</i>	<i>1 TB</i>
4	<i>VGA</i>	<i>NVIDIA GeForce 930mx</i>
5	<i>Keyboard</i>	Standar
6	<i>Mouse</i>	Standar

Tabel 2. *Software* yang digunakan

No	Spesifikasi	Keterangan
1	Sistem Operasi	<i>Windows 10 Pro</i>
2	Aplikasi Lembar Kerja	<i>Microsoft Excel</i>
3	Aplikasi Data <i>Mining</i>	<i>RapidMiner</i>

3.3 Jenis Penelitian

Jenis penelitian menggunakan penelitian kuantitatif dengan pendekatan deskriptif yaitu studi kasus pada toko Pusat Ponsel & Laptop Palangka Raya. Peneliti menggunakan jenis penelitian ini karena menggunakan angka, mulai dari pengumpulan data, penafsiran terhadap data tersebut, serta hasil yang sudah dianalisis akan dijabarkan secara deskriptif. Dalam penelitian ini, peneliti hanya berfokus pada satu permasalahan saja yaitu memprediksi penjualan *smartphone* berdasarkan data penjualan *smartphone* yang ada di toko Pusat Ponsel & Laptop Palangka Raya.

3.4 Teknik Pengumpulan Data

Teknik pengumpulan data yang digunakan sebagai berikut:

a. Observasi

Observasi merupakan suatu metode pengumpulan data dengan cara pengamatan langsung ke lapangan pada objek yang akan diteliti (Hardani, 2020).

Langkah awal, peneliti akan melakukan observasi langsung ke lokasi penelitian di Pusat Ponsel & Laptop Palangka Raya. Kegiatan observasi meliputi kondisi tempat penelitian serta kegiatannya.

b. Dokumentasi

Dokumentasi merupakan suatu metode pengumpulan data dengan melihat atau menganalisis dokumen-dokumen yang dibuat oleh subjek sendiri atau oleh orang lain tentang subjek (Hardani, 2020).

Sebagai langkah awal, peneliti akan bertanya dan meminta izin kepada pihak Pusat Ponsel & Laptop Palangka Raya untuk meminta data yang diperlukan seperti data – data penjualan *smartphone* dalam bentuk file *excel*.

c. Studi Kepustakaan

Studi kepustakaan merupakan suatu metode dalam mempelajari berbagai literatur yang relevan dari buku, artikel ilmiah, berita, maupun dari internet yang terkait dengan topik penelitian (Hardani, 2020).

Sebagai langkah awal, peneliti akan mempelajari tentang data *mining* dan *rapidminer* dari literatur yang ada dari buku, jurnal penelitian dan internet.

3.5 Analisis

a. Analisis Data

Data yang digunakan yaitu data penjualan *smartphone* di toko PPL Palangka Raya selama 6 bulan dari bulan Oktober 2021 sampai Maret 2022.

b. Analisis Proses

Data akan dianalisis dengan mengikuti proses tahapan dalam *Knowledge Discovery in Database (KDD)*, untuk menentukan informasi yang berguna serta pola-pola yang ada di dalam data. berikut tahapan- tahapannya:

1) Data Seleksi (*Selection*)

Data yang digunakan dalam penelitian ini berasal dari Pusat Ponsel & Laptop Palangka Raya, yaitu data penjualan *smartphone* pada bulan Oktober 2021 sampai Maret 2022. Semua atribut yang ada pada data penjualan diantaranya Nama *Item*, Kode *Item*, *Barcode*, Jumlah Terjual, Harga Beli, Harga Jual. Berdasarkan Tabel 3 – Tabel 8 dengan total penjualan sebanyak 1519 unit *smartphone*, atribut data akan diseleksi dan yang digunakan hanya 2 atribut yaitu Nama *Item* dan Jumlah Terjual untuk setiap bulan perminggunya.

Tabel 3. Data Penjualan *Smartphone* Oktober 2021

No	Nama Item	Kode Item	Barcode	Jumlah Terjual	Beli	Jual
1	OPPO A15 2/32 GB 866200055089899	001850	866200055089899	1	1,560,000	1,699,000
2	SAMSUNG A12 4/128 GB 350471513177807	18101	350471513177807	1	2,255,060	2,399,000
3	SAMSUNG A12 4/128 GB 350471513211762	18102	350471513211762	1	2,255,060	2,399,000
4	VIVO Y53S 11/128 GB	17981	868598058129358	1	3,330,000	3,700,000
5	VIVO Y53S 11/128 GB 868598058132170	17982	868598058132170	1	3,330,000	3,700,000

No	Nama Item	Kode Item	Barcode	Jumlah Terjual	Beli	Jual
...
186	SAMSUNG A03S 4/64GB BLACK IMEI : 3569775113595 55	CH73 YUT	3569775113 59555	1	1,673,070	1,799,000

Tabel 4. Data Penjualan *Smartphone* November 2021

No	Nama Item	Kode Item	Barcode	Jumlah Terjual	Beli	Jual
1	DEMO VIVO Y53S 11/128 GB 8685980595507 76	00179 9	8685980595 50776	1	2,230,000	2,300,000
2	OPPO A16 4/64 GB PERAK ANGKASA IMEI : 8671240550393 90	AHC TPYF	8671240550 39390	1	2,250,000	2,499,000
3	OPPO A16 3/32GB PERAK ANGKASA IMEI : 8664710506640 32	M7V THK A	8664710506 64032	1	1,800,000	1,999,000
4	OPPO A16 3/32GB PERAK ANGKASA IMEI : 8652450596533 32	U7YF AMK	8652450596 53332	1	1,800,000	1,999,000
5	OPPO A16 3/32GB PERAK ANGKASA IMEI : 8652450597117 91	X7K MNU L	8652450597 11791	1	1,800,000	1,999,000
...
247	REALME C25Y 4/64 GB ABU	KHR VEJ9	8601390525 13491	1	1,840,000	1,899,000

Tabel 5. Data Penjualan *Smartphone* Desember 2021

No	Nama Item	Kode Item	Barcode	Jumlah Terjual	Beli	Jual
1	DEMO OPPO RENO 5F 8/128 BLACK	10141 4	8657200501 11159	1	2,150,000	2,300,000
2	OPPO A16 4/64GB BLUE IMEI : 8666530544841 90	JVCM A9U	8666530544 84190	1	2,250,000	2,499,000
3	VIVO Y21S 4/128GB PEARL WHITE	MYV LEC4	8621940560 35593	1	2,520,000	2,800,000
4	OPPO A95 8/128GB HITAM BINTANG BERSINAR IMEI : 8626190523832 33	UHV7 EXP	8626190523 83233	1	3,600,000	3,999,000
5	OPPO A16 3/32GB PERAK ANGKASA IMEI : 8652450597118 58	CYW LMT4	8652450597 11858	1	1,800,000	1,999,000
...
250	SAMSUNG A12 6/128GB BLACK IMEI : 3546687765461 66	MFY AWTJ	3546687765 46166	1	2,324,070	2,499,000

Tabel 6. Data Penjualan *Smartphone* Januari 2022

No	Nama Item	Kode Item	Barcode	Jumlah Terjual	Beli	Jual
1	DEMOLIVE VIVO Y12s 3/32 BLUE 8680610596115 77	HP021 88405 52025 2448		1	1,250,000	1,400,000
2	OPPO A16 4/64GB SILVER	PYCK NT7	8671240575 65533	1	2,250,000	2,500,000

No	Nama Item	Kode Item	Barcode	Jumlah Terjual	Beli	Jual
250	REALME C11 2021 2/32GB LAKE BLUE IMEI : 8654620567332 11	KALC FEX	8654620567 33211	1	1,221,000	1,299,000

Tabel 8. Data Penjualan *Smartphone* Maret 2022

No	Nama Item	Kode Item	Barcode	Jumlah Terjual	Beli	Jual
1	OPPO A16 3/32 GB	17894	8652450548 30810	1	1,800,000	2,000,000
2	OPPO A16 3/32 GB 8652450563717 97	17896	8652450563 71797	1	1,800,000	1,999,000
3	OPPO A16 3/32 GB 8652450529444 31	17891	8652450529 44431	1	1,800,000	1,999,000
4	OPPO A16 3/32 GB 8652450531297 92	17892	8652450531 29792	1	1,800,000	1,999,000
5	OPPO A16 3/32 GB 8652450546236 11	17893	8652450546 23611	1	1,800,000	2,000,000
...
280	VIVO Y1S 2/32GB BLUE IMEI : 8694150591511 75	LFA3 P7K	8694150591 51175	1	1,449,000	1,449,000

2) *Preprocessing*

Pada tahap *preprocessing* dilakukan dua proses yaitu, integritas data (proses penggabungan data) dan data *cleaning* (proses membersihkan data yang tidak relevan, *missing value*, dan duplikat). Data penjualan akan digabungkan tiap bulan perminggunya untuk masing-masing *brand smartphone* serta penghapusan data penjualan yang tidak digunakan dan

duplikat sehingga total penjualan yang digunakan sebanyak 1511 unit *smartphone*. Hasil dari proses *preprocessing* dapat di lihat pada Tabel 9 – Tabel 13 untuk masing – masing *brand smartphone* dengan total penjualan untuk setiap *brand smartphone* seperti oppo sebanyak 394 unit, vivo sebanyak 726 unit, samsung sebanyak 156 unit, xiaomi sebanyak 120 unit, dan realme sebanyak 115 unit *smartphone*.

Tabel 9. Sampel Data Penjualan *Smartphone* Oppo

Bulan	Minggu			
	Ke-1	Ke-2	Ke-3	Ke-4
Oktober	6	12	11	7
November	14	8	20	17
Desember	18	16	22	17
Januari	35	33	19	23
Februari	13	19	9	13
Maret	15	4	23	20

Tabel 10. Sampel Data Penjualan *Smartphone* Vivo

Bulan	Minggu			
	Ke-1	Ke-2	Ke-3	Ke-4
Oktober	30	38	19	34
November	26	23	35	20
Desember	24	33	25	22
Januari	29	22	31	37
Februari	30	21	45	27
Maret	40	31	46	38

Tabel 11. Sampel Data Penjualan *Smartphone* Samsung

Bulan	Minggu			
	Ke-1	Ke-2	Ke-3	Ke-4
Oktober	1	3	0	2
November	8	13	14	11
Desember	6	5	3	13
Januari	6	4	2	13
Februari	10	7	7	8
Maret	5	9	1	5

Tabel 12. Sampel Data Penjualan *Smartphone* Xiaomi

Bulan	Minggu			
	Ke-1	Ke-2	Ke-3	Ke-4
Oktober	3	2	4	6
November	5	1	2	8
Desember	4	1	1	6
Januari	6	2	8	9
Februari	4	10	5	10
Maret	8	5	5	5

Tabel 13. Sampel Data Penjualan *Smartphone* Realme

Bulan	Minggu			
	Ke-1	Ke-2	Ke-3	Ke-4
Oktober	0	3	1	3
November	3	4	9	6
Desember	6	8	8	10
Januari	10	2	7	8
Februari	1	3	2	3
Maret	4	2	6	6

3) Data Transformasi (*Transformation*)

Tahapan selanjutnya yaitu hasil dari *preprocessing* akan ditransformasi menjadi data *training* dan data *testing*. Data *training* berdasarkan data yang ada, dan harus dipilih terlebih dahulu bagian atribut yang mempengaruhi penjualan *smartphone*, yang disebut dengan data target, data target akan menjadi atribut yang mendukung dalam proses data *mining*. Berikut salah satu contoh hasil dari data *transformation* untuk data *training* oppo pada Tabel 14.

Tabel 14. Data Training Penjualan *Smartphone* Oppo

No	Data 1	Data 2	Data 3	Target
1	6	12	11	7
2	12	11	7	14
3	11	7	14	8
4	7	14	8	20
5	14	8	20	17

No	Data 1	Data 2	Data 3	Target
6	8	20	17	18
7	20	17	18	16
8	17	18	16	22
9	18	16	22	17
10	16	22	17	35
11	22	17	35	33
12	17	35	33	19
13	35	33	19	23

Data *training* dikelompokkan menjadi 2 yaitu data *input* dan data target. Data *input* merupakan data 1 – data 3 yang berdasarkan data sebelum target. Data target diambil berdasarkan minggu ke-4 bulan Oktober 2021 dan data target selanjutnya dari minggu ke-1 bulan November 2021, dan dilakukan terus menerus sampai data target mencapai minggu ke-4 bulan Januari 2022. Berikut merupakan salah satu contoh untuk data *testing* (uji) untuk penjualan *smartphone* oppo pada Tabel 15.

Tabel 15. Data *Testing* Penjualan *Smartphone* Oppo

No	Data 1	Data 2	Data 3	Target
1	33	19	23	13
2	19	23	13	19
3	23	13	19	9
4	13	19	9	13
5	19	9	13	15
6	9	13	15	4
7	13	15	4	23
8	15	4	23	20

Data *testing* dibuat dengan cara yang sama dengan data *training* dan data yang digunakan yaitu data dari 2 bulan terakhir pada bulan Februari 2022 dan Maret 2022. Untuk data *training* dan data *testing* brand

smartphone lainnya dilakukan dengan cara yang sama. Berikut Tabel 16

– Tabel 23 merupakan data transformasi untuk *brand smartphone* lainnya.

Tabel 16. Data *Training* Penjualan *Smartphone* Vivo

No	Data 1	Data 2	Data 3	Target
1	30	38	19	34
2	38	19	34	26
3	19	34	26	23
4	34	26	23	35
5	26	23	35	20
6	23	35	20	24
7	35	20	24	33
8	20	24	33	25
9	24	33	25	22
10	33	25	22	29
11	25	22	29	22
12	22	29	22	31
13	29	22	31	37

Tabel 17. Data *Testing* Penjualan *Smartphone* Vivo

No	Data 1	Data 2	Data 3	Target
1	22	31	37	30
2	31	37	30	21
3	37	30	21	45
4	30	21	45	27
5	21	45	27	40
6	45	27	40	31
7	27	40	31	46
8	40	31	46	38

Tabel 18. Data *Training* Penjualan *Smartphone* Samsung

No	Data 1	Data 2	Data 3	Target
1	1	3	0	2
2	3	0	2	8
3	0	2	8	13
4	2	13	13	14
5	8	14	14	11
6	13	11	11	6
7	14	6	6	5
8	11	5	5	3
9	6	3	3	13
10	5	13	13	6

No	Data 1	Data 2	Data 3	Target
11	3	6	6	4
12	13	4	4	2
13	6	2	2	13

Tabel 19. Data *Testing* Penjualan *Smartphone* Samsung

No	Data 1	Data 2	Data 3	Target
1	4	2	13	10
2	2	13	10	7
3	13	10	7	7
4	10	7	7	8
5	7	7	8	5
6	7	9	5	9
7	8	5	9	1
8	5	9	1	5

Tabel 20. Data *Training* Penjualan *Smartphone* Xiaomi

No	Data 1	Data 2	Data 3	Target
1	3	2	4	6
2	2	4	6	5
3	4	6	5	1
4	6	5	1	2
5	5	1	2	8
6	1	2	8	4
7	2	8	4	1
8	8	4	1	1
9	4	1	1	6
10	1	1	6	6
11	1	6	6	2
12	6	6	2	8
13	6	2	8	9

Tabel 21. Data *Testing* Penjualan *Smartphone* Xiaomi

No	Data 1	Data 2	Data 3	Target
1	2	8	9	4
2	8	9	4	10
3	9	4	10	5
4	4	10	5	10
5	10	5	10	8
6	5	10	8	5
7	10	8	5	5
8	8	5	5	5

Tabel 22. Data *Training* Penjualan *Smartphone* Realme

No	Data 1	Data 2	Data 3	Target
1	0	3	1	3
2	3	1	3	3
3	1	3	3	4
4	3	3	4	9
5	3	4	9	6
6	4	6	6	6
7	6	6	6	8
8	6	6	8	8
9	6	8	8	10
10	8	8	10	10
11	8	10	10	2
12	10	10	2	7
13	10	2	7	8

Tabel 23. Data *Testing* Penjualan *Smartphone* Realme

No	Data 1	Data 2	Data 3	Target
1	2	7	8	1
2	7	8	1	3
3	8	1	3	2
4	1	3	2	3
5	3	2	3	4
6	2	3	4	2
7	3	4	2	6
8	4	2	6	6

4) Data *Mining*

Pada tahap ini, data yang sudah siap diolah akan dilakukan proses data *mining* menggunakan algoritma *K-Nearest Neighbor*. Contohnya yaitu menggunakan data *training* oppo pada Tabel 14 dan data *testing* dari penjualan *smartphone* Oppo untuk data no. 1 pada Tabel 15. Berikut langkah-langkah yang harus dilakukan:

- a) Menentukan nilai k. Penentuan nilai k umumnya minimal 1 sampai batas data *training* dikurang satu, berdasarkan jumlah data *training*

pada Tabel 14, nilai k yang bisa digunakan dari 1 – 12. Dalam penelitian ini nilai k yang digunakan yaitu 3.

- b) Melakukan perhitungan jarak data *training* dan data *testing* penjualan *smartphone* Oppo no. 1 menggunakan rumus jarak *Euclidean Distance* sebagai berikut :

$$\begin{aligned}d_1 &= \sqrt{(6 - 33)^2 + (12 - 19)^2 + (11 - 23)^2 + (7 - 13)^2} \\ &= 30,95\end{aligned}$$

$$\begin{aligned}d_2 &= \sqrt{(12 - 33)^2 + (11 - 19)^2 + (7 - 23)^2 + (14 - 13)^2} \\ &= 27,60\end{aligned}$$

$$\begin{aligned}d_3 &= \sqrt{(11 - 33)^2 + (7 - 19)^2 + (14 - 23)^2 + (8 - 13)^2} \\ &= 27,09\end{aligned}$$

$$\begin{aligned}d_4 &= \sqrt{(7 - 33)^2 + (14 - 19)^2 + (8 - 23)^2 + (20 - 13)^2} \\ &= 31,22\end{aligned}$$

$$\begin{aligned}d_5 &= \sqrt{(14 - 33)^2 + (8 - 19)^2 + (20 - 23)^2 + (17 - 13)^2} \\ &= 22,52\end{aligned}$$

$$\begin{aligned}d_6 &= \sqrt{(8 - 33)^2 + (20 - 19)^2 + (17 - 23)^2 + (18 - 13)^2} \\ &= 26,21\end{aligned}$$

$$\begin{aligned}d_7 &= \sqrt{(20 - 33)^2 + (17 - 19)^2 + (18 - 23)^2 + (16 - 13)^2} \\ &= 14,39\end{aligned}$$

$$\begin{aligned}d_8 &= \sqrt{(17 - 33)^2 + (18 - 19)^2 + (16 - 23)^2 + (22 - 13)^2} \\ &= 19,67\end{aligned}$$

$$\begin{aligned}d_9 &= \sqrt{(18 - 33)^2 + (16 - 19)^2 + (22 - 23)^2 + (17 - 13)^2} \\ &= 15,84\end{aligned}$$

$$d_{10} = \sqrt{(16 - 33)^2 + (22 - 19)^2 + (17 - 23)^2 + (35 - 13)^2}$$

$$= 28,60$$

$$d_{11} = \sqrt{(22 - 33)^2 + (17 - 19)^2 + (35 - 23)^2 + (33 - 13)^2}$$

$$= 25,87$$

$$d_{12} = \sqrt{(17 - 33)^2 + (35 - 19)^2 + (33 - 23)^2 + (19 - 13)^2}$$

$$= 25,46$$

$$d_{13} = \sqrt{(35 - 33)^2 + (33 - 19)^2 + (19 - 23)^2 + (23 - 13)^2}$$

$$= 17,78$$

- c) Melakukan pengurutan jarak hasil perhitungan dari *Euclidean Distance* dari terkecil sampai terbesar.

Tabel 24. Pengurutan hasil *Euclidean Distance*

No	Urutan	<i>Euclidean Distance</i>
1	d7	14,39
2	d9	15,84
3	d13	17,78
4	d8	19,67
5	d5	22,52
6	d12	25,46
7	d11	25,87
8	d6	26,21
9	d3	27,09
10	d2	27,60
11	d10	28,60
12	d1	30,95
13	d4	31,22

- d) Menentukan kelompok data hasil uji berdasarkan target mayoritas dari k tetangga terdekat yaitu 3. Oleh karena itu diambil 3 jarak terkecil yaitu d7, d9, dan d13. Dikarenakan data targetnya angka maka dilakukan perhitungan nilai rata-rata dari atribut target pada data

training di Tabel 14 dan dibagi dengan nilai k yang sudah ditentukan, yaitu: Data uji 1 = $(16 + 17 + 23) / 3 = 18,6$

- e) Hasil prediksi penjualan *smartphone* Oppo untuk minggu ke-1 pada bulan April 2022 sebanyak 18 unit. Proses perhitungannya akan terus dilakukan dengan cara yang sama sampai data *testing* (uji) terakhir.

5) *Interpretation/Evaluation*

Tujuan akhir dari seluruh tahapan yang ada yaitu penyajian informasi, berupa pola informasi hasil dari proses data *mining* menggunakan metode *K-Nearest Neighbor*. Selanjutnya melakukan evaluasi menggunakan *Root Mean Squared Error (RMSE)* untuk mengetahui keakuratan kinerja model suatu algoritma, semakin kecil nilai *error* yang dihasilkan maka semakin dekat nilai atau jarak antara nilai aktual dengan nilai prediksi. Berikut rumus *Root Mean Squared Error (RMSE)* pada persamaan 2:

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

Y' = Nilai Prediksi

Y = Nilai Target

n = Jumlah Data

BAB IV

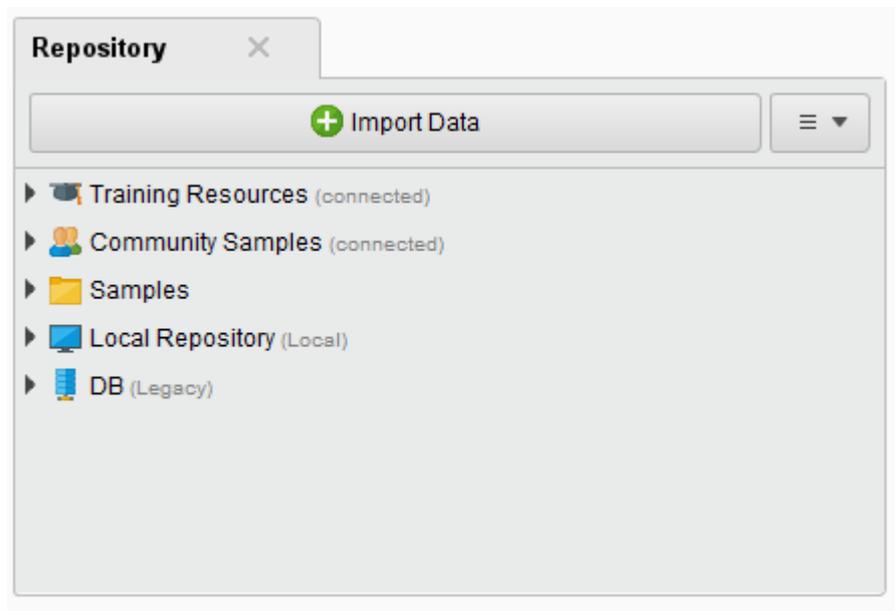
HASIL DAN PEMBAHASAN

4.1 Pengujian Algoritma

Pengujian yang dilakukan penulis ini menggunakan metode *K-Nearest Neighbor* dan aplikasi *RapidMiner*. Berikut tahapan yang dilakukan dalam pengujian algoritma dengan menggunakan *RapidMiner* :

a. Persiapan data

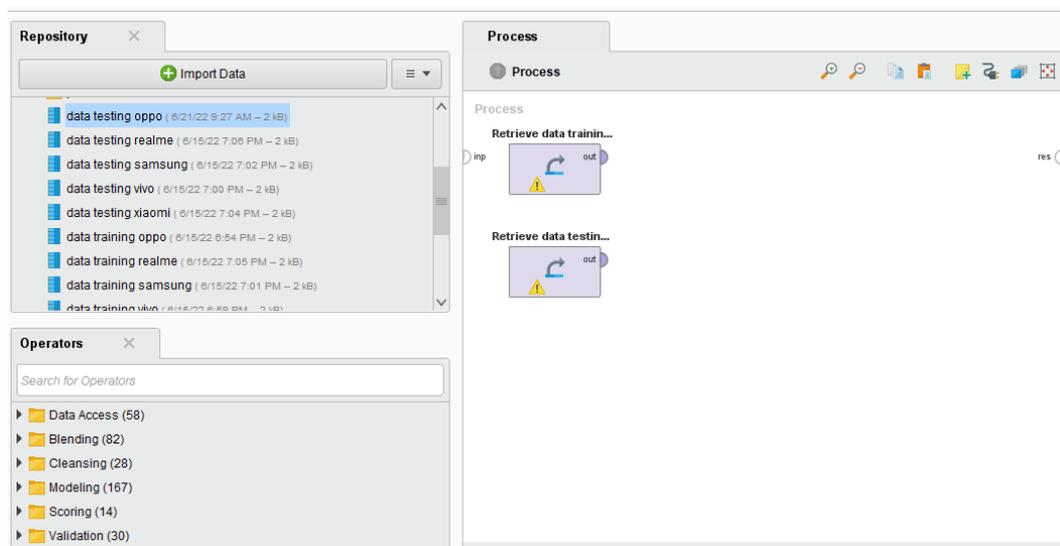
Data *training* dan data *testing* yang sudah dibuat harus dimasukkan terlebih dahulu kedalam aplikasi *RapidMiner* dengan cara pilih “*Import Data*” pada tampilan *RapidMiner* seperti pada Gambar 4.



Gambar 4. *Import Data*

b. Memasukkan data yang digunakan

Selanjutnya, data *training* dan data *testing* yang sudah dimasukkan akan di *drag* ke “*Process*” seperti pada Gambar 5.

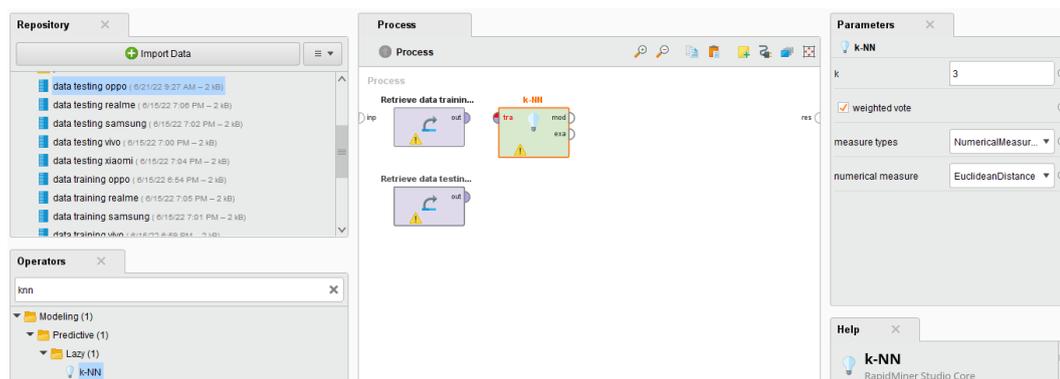


Gambar 5. Data *Training* dan Data *Testing* di masukkan kedalam *Process*

Berdasarkan Gambar 5, contoh data yang digunakan yaitu data *training* oppo pada Tabel 14 dan data *testing* oppo pada Tabel 15.

c. Pemilihan algoritma

Setelah data sudah dimasukkan ke dalam “*Process*”, lalu langkah selanjutnya memasukkan algoritma yang akan digunakan. Penulis akan memilih algoritma *K-Nearest Neighbor* pada bagian “*Operators*”. Setelah dipilih, lalu lakukan *drag* atau tekan 2x untuk masuk ke dalam “*process*” seperti pada Gambar 6.



Gambar 6. Operator K-NN

Berdasarkan Gambar 6, algoritma *K-NN* diatur terlebih dahulu pada bagian “*Parameters*” perhitungan yang digunakan yaitu *Euclidean Distance* dan *K* yang digunakan yaitu 3.

d. Menentukan operator

Operator diperlukan untuk menghitung dan menampilkan hasil proses perhitungannya. Operator yang digunakan oleh penulis adalah “*Apply Model*” dan “*Performance*” seperti pada Gambar 7.



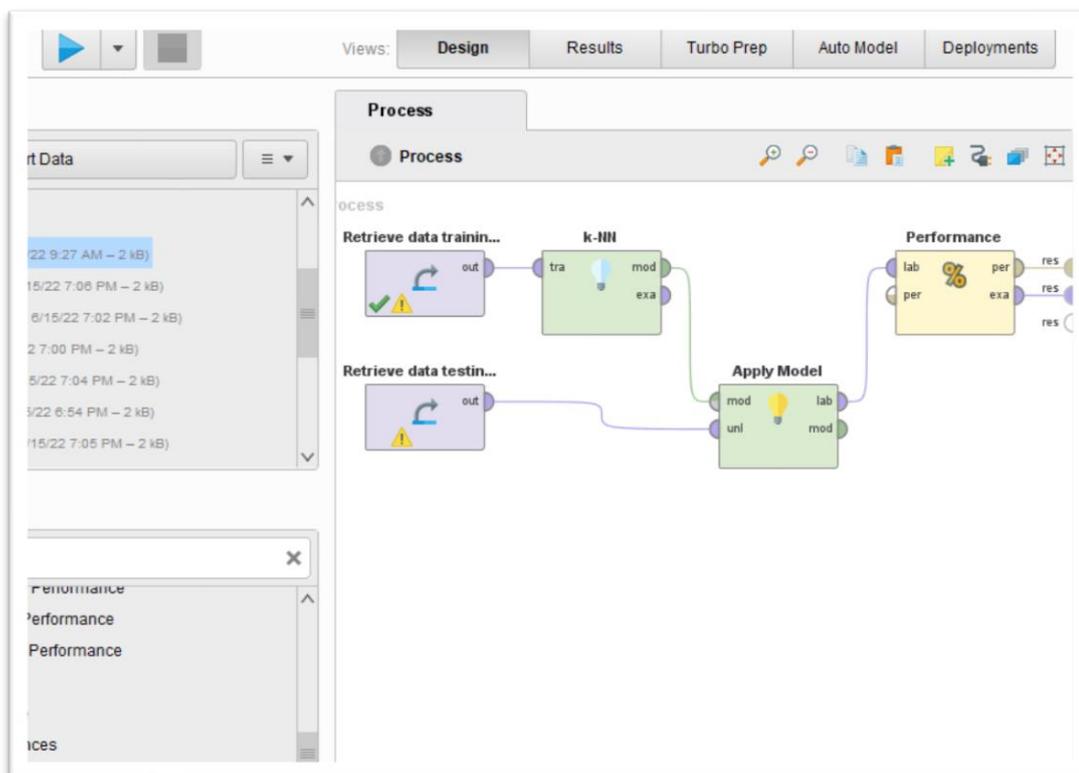
Gambar 7. Operator *Apply Model* dan *Performance*

Berdasarkan Gambar 7, operator “*Apply Model*” digunakan untuk menerapkan hasil yang telah dilatih menggunakan *K-NN* pada data

training kepada data *testing*. Operator “*Performance*” digunakan untuk melakukan evaluasi algoritma yang digunakan dan menampilkan hasil prediksi penjualannya.

e. Menghubungkan data yang digunakan dengan operator

Setelah data dan operator yang akan digunakan sudah masuk kedalam “*Process*”, yang harus dilakukan selanjutnya adalah menghubungkannya seperti pada Gambar 8, karena proses tidak bisa berjalan jika belum disambungkan.



Gambar 8. Susunan Operator

Berdasarkan Gambar 8, setelah semuanya sudah disambungkan, maka proses dapat dimulai dengan menekan tombol “F11” pada *keyboard*.

f. Hasil

Berikut hasil prediksi penjualan untuk *brand smartphone* oppo dapat dilihat pada Gambar 9.

Row No.	Data 1	Data 2	Data 3	Target	prediction(T...
1	33	19	23	13	18.660
2	19	23	13	19	25.002
3	23	13	19	9	18.083
4	13	19	9	13	18.691
5	19	9	13	15	12.952
6	9	13	15	4	11.296
7	13	15	4	23	14.132
8	15	4	23	20	14.166

Gambar 9. Hasil Prediksi Penjualan *Smartphone* Oppo

Berdasarkan Gambar 9, merupakan hasil perhitungan dari data *training* oppo pada Tabel 14 yang telah dilatih pada data *testing* oppo pada Tabel 15 dengan jumlah data *testing* sebanyak 8, diperoleh hasil prediksinya untuk 2 bulan yang akan datang untuk setiap minggunya. Selanjutnya akan di totalkan hasil prediksinya dan dibulatkan nilainya. Hasil total prediksi oppo = $18 + 25 + 18 + 18 + 12 + 11 + 14 + 14 = 130$ unit *smartphone*.

Untuk hasil prediksi penjualan *brand smartphone* lainnya dilakukan dengan cara yang sama. Berikut hasil prediksi untuk *brand smartphone* lainnya pada Gambar 10 – Gambar 13.

Row No.	Data 1	Data 2	Data 3	Target	prediction(T...
1	22	31	37	30	22.686
2	31	37	30	21	26.629
3	37	30	21	45	32.337
4	30	21	45	27	27.352
5	21	45	27	40	23.002
6	45	27	40	31	27.632
7	27	40	31	46	26.082
8	40	31	46	38	27.538

Gambar 10. Hasil Prediksi Penjualan *Smartphone Vivo*

Berdasarkan Gambar 10, hasil total prediksi untuk *smartphone vivo* sebagai berikut: Vivo = 22 + 26 + 32 + 27 + 23 + 27 + 26 + 27 = 210 unit *smartphone*.

Row No.	Data 1	Data 2	Data 3	Target	prediction(T...
1	4	2	13	10	9.976
2	2	13	10	7	10.420
3	13	10	7	7	4.761
4	10	7	7	8	3.368
5	7	7	8	5	6.391
6	7	9	5	9	6.486
7	8	5	9	1	6.442
8	5	9	1	5	9.907

Gambar 11. Hasil Prediksi Penjualan *Smartphone Samsung*

Berdasarkan Gambar 11, hasil total prediksi untuk *smartphone Samsung* sebagai berikut: Samsung = 9 + 10 + 4 + 3 + 6 + 6 + 6 + 9 = 53 unit *smartphone*.

Row No.	Data 1	Data 2	Data 3	Target	prediction(T...
1	2	8	9	4	1.363
2	8	9	4	10	3.828
3	9	4	10	5	5.330
4	4	10	5	10	1.289
5	10	5	10	8	6.117
6	5	10	8	5	1.318
7	10	8	5	5	3.436
8	8	5	5	5	3.408

Gambar 12. Hasil Prediksi Penjualan *Smartphone* Xiaomi

Berdasarkan Gambar 12, hasil total prediksi untuk *smartphone* xiaomi sebagai berikut: $\text{Xiaomi} = 1 + 3 + 5 + 1 + 6 + 1 + 3 + 3 = 23$ unit *smartphone*.

Row No.	Data 1	Data 2	Data 3	Target	prediction(T...
1	2	7	8	1	6.605
2	7	8	1	3	7.022
3	8	1	3	2	6.648
4	1	3	2	3	4.786
5	3	2	3	4	5.347
6	2	3	4	2	5.738
7	3	4	2	6	5.489
8	4	2	6	6	6.137

Gambar 13. Hasil Prediksi Penjualan *Smartphone* Realme

Berdasarkan Gambar 13, hasil total prediksi untuk *smartphone* realme sebagai berikut: $\text{Realme} = 6 + 7 + 6 + 4 + 5 + 5 + 5 + 6 = 44$ unit *smartphone*.

4.2 Evaluasi Algoritma

Evaluasi perhitungan data akan menggunakan algoritma yang sudah dihitung otomatis oleh aplikasi *RapidMiner* berupa *Root Mean Square Error (RMSE)* yaitu penjumlahan kuadrat *error* atau selisih antara nilai sebenarnya dan nilai prediksi, kemudian membagi jumlah tersebut dengan banyaknya jumlah data *testing* dan menghitung akarnya. Hasil nilai *RMSE* yang mendekati angka nol, maka semakin akurat nilai yang diprediksi mendekati dengan nilai aslinya. Berikut nilai *RMSE* untuk semua *brand smartphone* pada Tabel 25.

Tabel 25. Nilai *RMSE* Untuk Semua *Brand Smartphone*

No	Brand	Nilai RMSE K = 3
1	Oppo	6,643
2	Vivo	11,468
3	Samsung	3,538
4	Xiaomi	4,23
5	Realme	3,325

Berdasarkan Tabel 25, dapat disimpulkan bahwa nilai *RMSE* untuk masing-masing *brand smartphone* dari yang terkecil sampai terbesar, yaitu realme dengan nilai 3,325, samsung dengan nilai 3,538, xiaomi dengan nilai 4,23, oppo dengan nilai 6,643, dan vivo dengan nilai 11,468. Perhitungan untuk mendapatkan nilai *RMSE* dilakukan dengan contoh menggunakan *brand smartphone* oppo pada Gambar 9, bagian target dan prediksi dengan jumlah *n* sebanyak 8. Untuk *brand smartphone* lainnya dihitung dengan cara yang sama.

RMSE Oppo :

$$\sqrt{\frac{(18,65 - 13)^2 + (25 - 19)^2 + (18,08 - 9)^2 + (18,69 - 13)^2 + (12,95 - 15)^2 + (11,29 - 4)^2 + (14,13 - 23)^2 + (14,16 - 20)^2}{8}}$$

$$= 6,643$$

RMSE Vivo :

$$\sqrt{\frac{(22,68 - 30)^2 + (26,62 - 21)^2 + (32,33 - 45)^2 + (27,35 - 27)^2 + (23 - 40)^2 + (27,63 - 31)^2 + (26,08 - 46)^2 + (27,53 - 38)^2}{8}}$$

$$= 11,468$$

RMSE Samsung :

$$\sqrt{\frac{(9,97 - 10)^2 + (10,42 - 7)^2 + (4,76 - 7)^2 + (3,36 - 8)^2 + (6,39 - 5)^2 + (6,48 - 9)^2 + (6,44 - 1)^2 + (9,90 - 5)^2}{8}}$$

$$= 3,538$$

RMSE Xiaomi :

$$\sqrt{\frac{(1,36 - 4)^2 + (3,82 - 10)^2 + (5,33 - 5)^2 + (1,289 - 10)^2 + (6,11 - 8)^2 + (1,31 - 5)^2 + (3,43 - 5)^2 + (3,40 - 5)^2}{8}}$$

$$= 4,23$$

RMSE Realme :

$$\sqrt{\frac{(6,60 - 1)^2 + (7,02 - 3)^2 + (6,64 - 2)^2 + (4,78 - 3)^2 + (5,34 - 4)^2 + (5,73 - 2)^2 + (5,48 - 6)^2 + (6,13 - 6)^2}{8}}$$

$$= 3,325$$

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil dan pembahasan algoritma pada penelitian ini menggunakan aplikasi *RapidMiner* dapat disimpulkan bahwa :

- a. Dengan menggunakan data penjualan *smartphone* selama 6 bulan dari Oktober 2021 – Maret 2022 berdasarkan tahapan *KDD* lalu diterapkan pada data *mining* menggunakan metode *K-Nearest Neighbor* dengan nilai $K=3$ menunjukkan bahwa algoritma dapat memprediksi jumlah penjualan *smartphone* untuk 2 bulan yang akan datang.
- b. Hasil prediksi untuk setiap brand *smartphone* yang paling laris adalah vivo sebanyak 210 unit, oppo sebanyak 130 unit, samsung sebanyak 53 unit, realme sebanyak 44 unit, dan xiaomi sebanyak 23 unit *smartphone*.
- c. Hasil nilai *Root Mean Squared Error (RMSE)* yang diperoleh oleh setiap *brand smartphone*, oppo dengan nilai 6,643, vivo dengan nilai 11,468, samsung dengan nilai 3,538, xiaomi dengan nilai 4,23, dan realme dengan nilai 3,325. Berdasarkan nilai *RMSE* pada tiap *brand smartphone* dapat disimpulkan bahwa *brand smartphone* realme memiliki *RMSE* terkecil, yang berarti memiliki hasil prediksi yang mendekati dengan nilai yang sebenarnya.

5.2 Saran

Penelitian ini memiliki beberapa kekurangan karena menggunakan aplikasi bantuan berupa *RapidMiner* dan diharapkan kepada penulis lain untuk dapat mengembangkannya lagi. Berikut saran yang dapat disampaikan untuk dapat mengembangkan penelitian ini :

- a. Penelitian selanjutnya diharapkan dapat dibandingkan menggunakan metode prediksi lainnya seperti algoritma *C4.5*, algoritma *k-means*, algoritma *apriori*, dan algoritma *naive bayes*.
- b. Penelitian selanjutnya diharapkan menggunakan aplikasi bantuan selain *RapidMiner* atau dibuat menjadi sebuah program menggunakan bahasa pemrograman seperti *PHP*, *Python*, *Java* atau *C++*.

DAFTAR PUSTAKA

- Amalia, Y. R., 2018. Penerapan Data *Mining* Untuk Prediksi Penjualan Produk Elektronik Terlaris Menggunakan Metode *K-Nearest Neighbor* (Studi Kasus: PT. Bintang Multi Sarana Palembang). *Program Studi Sistem Informasi Fakultas Sains Dan Teknologi Universitas Islam Negeri Raden Fatah Palembang*.
- Berman, B., & Evans, J. R., 2012. *Retail Management, A Strategic Approach. Edisi 11*. New Jersey: Pearson Education Limited.
- Dewi, S.P., Nurwati, N. & Rahayu, E., 2022. Penerapan Data *Mining* Untuk Prediksi Penjualan Produk Terlaris Menggunakan Metode *K-Nearest Neighbor*. *Building of Informatics, Technology and Science (BITS)*, 3(4), pp. 639-648.
- Hardani, 2020. *Metode Penelitian Kualitatif & Kuantitatif*. Pustaka Ilmu.
- Kadir A., 2014. *Pengenalan Sistem Informasi Edisi Revisi*. Yogyakarta: ANDI.
- Kaswandari, Y., 2017. Rancang Bangun Sistem Kendali Tirai *Vertical Blind* dan Lampu Ruangan Melalui *Smartphone Android* Berbasis *Arduino*. *Program Studi Teknik Telekomunikasi Politeknik Negeri Sriwijaya*.
- Kotler, P., & Keller, K. L., 2016. *Marketing Management, 15th Edition*. Harlow, Essex: Pearson Education Limited.
- Longkutoy, J. J., 2012. *Pengenalan Komputer*. Jakarta: Mutiara Sumber Widya.
- Putri, A. A., 2021. Penerapan Data *Mining* Untuk Memprediksi Penjualan Buah Dan Sayur Menggunakan Metode *K-Nearest Neighbor* (Studi Kasus: PT. Central Brastagi Utama). *Resolusi: Rekayasa Teknik Informatika dan Informasi*, 1(6), pp. 354-361.
- Rahmat, B. C. T. I., Gafar, A. A., Fajriani, N., Ramdani, U., Uyun, F. R., Purnamasari, Y., & Ransi, N., 2017. Implementasi *k-means clustering* pada *rapidminer* untuk analisis daerah rawan kecelakaan. *Seminar Nasional Riset Kuantitatif Terapan*, pp. 58-62.
- Rizqifaluthi, H. & Yaqin, M. A., 2018. *Process Mining Akademik Sekolah menggunakan RapidMiner*. *MATICS*, 10(2), pp. 47-51.
- Rerung, R. R., 2018. Penerapan Data *Mining* Dengan Memanfaatkan Metode *Association Rule* Untuk Promosi Produk. *J. Teknol. Rekayasa*, 3(1), p. 89.
- Suntoro, J., 2019. *Data Mining: Algoritma dan Implementasi dengan pemrograman PHP*. Jakarta: PT Elex Media Komputindo.
- Susanto, A., 2013. *Sistem Informasi Akuntansi*. Bandung: Lingga Jaya.
- Sutabri, T., 2012. *Analisis Sistem Informasi*. Yogyakarta: Andi.

- Witten, I. H., Frank, E., Hall, M. A., & Pal, C. J., 2016. *Data Mining: Practical machine learning tools and techniques (4th ed)*. Morgan Kaufmann Series in Data Management System.
- WPR, A.A., Rozi, F. & Sukmana, F., 2021. Prediksi Penjualan Produk Unilever Menggunakan Metode *K-Nearest Neighbor*. *JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, 6(1), pp. 155-160.
- Yanosma, D., Johar, A. & Anggriani, K., 2016. Implementasi Metode *K-Nearest Neighbor (KNN)* dan *Simple Additive Weighting (SAW)* dalam Pengambilan Keputusan Seleksi Anggota PASKIBRAKA. *Rekursif: Jurnal Informatika*, 4(2).
- Yuliana, Y., 2017. Pengaruh Pendidikan Asrama Dan Pembentukan Karakter Terhadap Hasil Belajar Sosiologi Siswa PPLP Sumatera Barat Kelas XI IPS SMAN 5 Padang. Program Studi Pendidikan Sosiologi STKIP PGRI Sumatera Barat.
- Yolanda, I. & Fahmi, H., 2020. Penerapan Data *Mining* Untuk Prediksi Penjualan Produk Roti Terlaris Pada PT. Nippon Indosari Corpindo Tbk Menggunakan Metode *K-Nearest Neighbor*. *Jurnal Ilmu Komputer dan Sistem Informasi (JIKOMSI)*, 3(1.1), pp. 9-15.

LAMPIRAN



**SEKOLAH TINGGI MANAJEMEN INFORMATIKA DAN KOMPUTER
(STMIK) PALANGKARAYA**

Jl. G. Obos No.114 Telp.0536-3224593, 3225515 Fax.0536-3225515 Palangka Raya
email : humas@stmikplk.ac.id - website : www.stmikplk.ac.id

SURAT TUGAS

No.99/STMIK-3.C.2/KP/III/2022

Ketua Program Studi Teknik Informatika Sekolah Tinggi Manajemen Informatika dan Komputer (STMIK) Palangkaraya menugaskan nama- nama tersebut di bawah ini :

1. Nama : Veny Cahya Hardita, M.Kom
NIK : 199504302020002
Sebagai Pembimbing I Dalam Pembuatan Program
2. Nama : Elok Faiqotul Himmah, S.Si., M.Sc.
NIK : 198503092009003
Sebagai Pembimbing II Dalam Penulisan Tugas Akhir

Untuk membimbing Tugas Akhir mahasiswa :

Nama : Braen Dwiatmajaya Garang
NIM : C1755201090
Program Studi : TEKNIK INFORMATIKA (55201)
Tanggal Daftar : 25 Februari 2022
Judul Tugas Akhir : Penerapan Data Mining Untuk Prediksi Penjualan Smartphone Paling Laris Menggunakan Metode K-Nearest Neighbor

Demikian surat ini dibuat agar dapat dipergunakan sebagaimana mestinya dan dilaksanakan dengan penuh tanggung jawab.

Palangka Raya, 16 Maret 2022
Ketua Program Studi Teknik Informatika,



Tembusan :

1. Pembimbing I dan II
2. Mahasiswa yang bersangkutan
3. Arsip



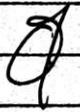
SEKOLAH TINGGI MANAJEMEN INFORMATIKA DAN KOMPUTER
(STMIK) PALANGKARAYA

Jl. G. Obos No.114 Telp.0536-3224593, 3225515 Fax.0536-3225515 Palangkaraya
email : humas@stmikpik.ac.id - webalta : www.stmikpik.ac.id

KARTU KEGIATAN KONSULTASI
TUGAS AKHIR

Nama Mahasiswa : Broen Dwiatmajaya Garang
NIM : C1755201090
Tanggal Persetujuan Judul : ~~25 Februari 2022~~ 25 Februari 2022
Judul Tugas Akhir : Penerapan Data Mining Untuk Prediksi Penjualan Smartphone Paling Laris Menggunakan Metode K-Nearest Neighbor

No.	Tanggal Konsultasi		Uraian	Tanda Tangan
	Terima	Kembali		
1	23/03/2022	24/03/2022	Jenis penelitian kualitatif kurang tepat sama topik	
2	03/04/2022	04/04/2022	latar belakang harus saling berhubungan; gunakan italic untuk tulisan bahasa asing, perbaiki penulisan yang kurang tepat	
3	06/04/2022	06/04/2022	Perbaiki kalimat pada perumusan masalah	
4	07/04/2022	08/04/2022	beri nomor persamaan untuk setiap persamaan mlk, tambahkan pembahasan tentang rapidminer, Embahkan kalimat pengantar pada perincian alat dan bahan	
5	08/04/2022	09/04/2022	Ceritakan apa yang kamu dapat dalam pengumpulan data tambahkan izin penelitian dan surat balasan penelitian	
6	08/04/2022	09/04/2022	tambahkan nomor judul pada tiap tabel, perkecil nomor dan urutan, tambahkan jadwal penelitian	
7	11/04/2022	12/04/2022	perbaiki penulisan sitasi, perbaiki jadwal penelitian, perbaiki jenis penelitian	
8	13/04/2022	14/04/2022	Acc Seminar	
9	13/04/2022	14/04/2022	Acc Seminar	

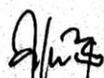
No.	Tanggal Konsultasi		Uraian	Tanda Tangan
	Terima	Kembali		
10	15/06/2022	16/06/2022	tambahkan metode yang dipakai dibatasi masalah jadwal penelitian tidak perlu lagi	
11	18/06/2022	20/06/2022	bab 2-4 disebutkan apa saja teorinya yang dijelaskan di bab 2, apa saja yang dijelaskan di bab 3 dan 4 secara umum tambahkan hasil evaluasi dengan RMSE di kesimpulan buat intrsafi	
12	21/06/2022	22/06/2022	perbaiki intrsafi paragraf isinya tentang latar belakang, secara singkat dan jelas, permasalahan, tujuan, metode dan hasil penelitian	
13	22/06/2022	22/06/2022	Acc Sidang	
14	22/06/2022	22/06/2022	Acc Sidang	

Menyetujui :

Dosen Pembimbing I,


YENY CAHYA H

Dosen Pembimbing II,


Elok Fatmahan Himmas



SEKOLAH TINGGI MANAJEMEN INFORMATIKA DAN KOMPUTER
STMIK PALANGKARAYA

Jl. G. Obos No. 114 ~ Telp. 0536-3224593 ~ Fax. 0536-3225515 Palangka Raya

Email: humas@stmikplk.ac.id ~ Website: www.stmikplk.ac.id

Nomor : 100.../STMIK-C.2.../AK.../III/2022
Lampiran : -
Perihal : Permohonan Izin Penelitian dan Pengumpulan Data untuk Tugas Akhir

Kepada
Yth. **Pimpinan CV. Media Grup Internasional**
Di -
Palangka Raya

Dengan hormat,

Sehubungan dengan penyusunan Tugas Akhir mahasiswa sebagai persyaratan kelulusan Program Studi Teknik Informatika (S1) pada STMIK Palangkaraya, maka dengan ini kami sampaikan permohonan izin penelitian dan pengumpulan data bagi mahasiswa kami berikut:

Nama : BRAEN DWIATMAJAYA GARANG
NIM : C1755201090
Prodi (Jenjang) : Teknik Informatika (S1)
Thn. Akad. (Semester) : 2021/2022 (10)
Lama Penelitian : 30 Maret 2022 s.d 30 April 2022
Tempat Penelitian : Pusat Ponsel dan Laptop

Dengan judul Tugas Akhir:

**PENERAPAN DATA MINING UNTUK PREDIKSI PENJUALAN
SMARTPHONE PALING LARIS MENGGUNAKAN METODE K-NEAREST
NEIGHBOR**

Adapun ketentuan dan aturan pemberian informasi dan data yang diperlukan dalam penelitian tersebut menyesuaikan dengan ketentuan/peraturan pada instansi Bapak/Ibu.

Demikian permohonan ini disampaikan, atas perhatian dan kerja samanya diucapkan terima kasih.

Palangka Raya, 30 Maret 2022
Ketua,

Suparno, M.Kom.
NIK. 196901041995105

PUSAT PONSEL & LAPTOP

Jln. Yos Sudarso no. 6J Kota Palangka Raya, Kalimantan Tengah 73312

Hal : Balasan

Kepada Yth :
Ketua STMIK Palangka Raya
Bapak Suparno, M.Kom
Di –
Palangka Raya

Dengan hormat,

Yang bertanda tangan di bawah ini :

Nama : Syahidin
Jabatan : Sales Head (Kepala Penjualan HP)

Menerangkan bahwa,

Nama : Braen Dwiatmajaya Garang
NIM : C1755201090
Jurusan : Teknik Informatika (S1)
Universitas : STMIK Palangka Raya

Telah kami setuju untuk mengadakan penelitian di toko Pusat Ponsel & Laptop dengan permasalahan dan judul :

**PENERAPAN DATA MINING UNTUK PREDIKSI PENJUALAN SMARTPHONE
PALING LARIS MENGGUNAKAN METODE K-NEAREST NEIGHBOR**

Demikian surat ini kami sampaikan, dan atas kerja samanya kami mengucapkan terima kasih.

Palangka Raya, 4 April 2022

Hormat Kami,

Sales Head Toko Pusat Ponsel & Laptop



Syahidin



SEKOLAH TINGGI MANAJEMEN INFORMATIKA DAN KOMPUTER
(STMIK) PALANGKARAYA

Jl. G. Obos No.114 Telp.0536-3224593, 3225515 Fax.0536-3225515 Palangka Raya
email : humas@stmikplk.ac.id - website : www.stmikplk.ac.id

SURAT TUGAS PENGUJI TUGAS AKHIR

No. 181/STMIK-3.C.2/KP/VI/2022

Ketua Program Studi Teknik Informatika Sekolah Tinggi Manajemen Informatika dan Komputer (STMIK) Palangkaraya menugaskan kepada nama- nama berikut :

1. Nama : Hotmian Sitohang, M.Kom.
NIK : 198503282008002
Sebagai Ketua
2. Nama : Ir. Hj. Siti Maryamah, M.M.
NIK : 196902021995004
Sebagai Sekretaris
3. Nama : Lili Rusdiana, M.Kom.
NIK : 198707282011007
Sebagai Anggota
4. Nama : Veny Cahya Hardita, M.Kom
NIK : 199504302020002
Sebagai Anggota
5. Nama : Elok Faiqotul Himmah, S.Si., M.Sc.
NIK : 198503092009003
Sebagai Anggota

Tim Penguji Tugas Akhir Mahasiswa :

- Nama : Braen Dwiatmajaya Garang
NIM : C1755201090
Hari/ Tanggal Ujian : Sabtu, 25 Juni 2022
Waktu : 13.00 WIB
Judul Tugas Akhir : Penerapan Data Mining Untuk Prediksi Penjualan Smartphone Paling Laris Menggunakan Metode K-Nearest Neighbor

Demikian surat ini dibuat agar dapat dipergunakan sebagaimana mestinya dan dilaksanakan dengan penuh tanggung jawab.

Palangka Raya, 23 Juni 2022
Ketua Program Studi Teknik Informatika,

Lili Rusdiana, M.Kom.
NIK : 198707282011007





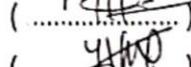
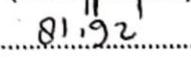
SEKOLAH TINGGI MANAJEMEN INFORMATIKA DAN KOMPUTER
(STMIK) PALANGKARAYA

Jl. G. Obos No.114 Telp.0536-3224593, 3225515 Fax.0536-3225515 Palangkaraya
email : humas@stmikpik.ac.id - website : www.stmikpik.ac.id

BERITA ACARA
UJIAN TUGAS AKHIR

Periode (Bulan) : Juni Tahun .. 2022

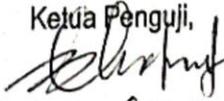
1. Hari/Tanggal Ujian : Sabtu / 25
2. Waktu (Jam) : 13.00 WIB sampai dengan WIB
3. Nama Mahasiswa : Braen Dwiatmaja Garang
4. Nomor Induk Mahasiswa : C.1755201090
5. Program Studi : Teknik Informatika
6. Tahun Angkatan : 2017
7. Judul Tugas Akhir : Pengolahan Data Mining Untuk Prediksi
Penjualan Smartphone Paling Laris
Menggunakan Metode K-Nearest Neighbor
8. Dosen Penguji :

Nama	Nilai	Tanda Tangan
1. Hatmian Sitohang, M.Kom =	2	()
2. Intji Siti Maryamah, M.M =		()
3. Abdul Hadi, S.T. Lili Rusdiana, M.Kom =		()
4. Vany Cahya Hordita, M.Kom =		()
5. Erok Foigat Himmeh, S.Si., M.Sc =		()
9. Hasil Ujian : LULUS / ~~TIDAK LULUS~~ *) NILAI = 81,92
10. Catatan Penting : 1. Lama Perbaikan : 10 hari
2. Jika lebih dari 1 (satu) bulan dikenakan sanksi berupa denda sebesar Rp. 600.000,- (Enam ratus ribu rupiah) per bulan dari tanggal ujian
3. Jika lebih dari 3 (tiga) bulan dari tanggal ujian maka hasil ujian dibatalkan dan wajib mengajukan judul dan pembimbing baru

Palangka Raya, 25 Juni 2022

Mengetahui :
Ketua Program Studi Teknik Informatika,


Lili Rusdiana, M.Kom,
NIK. 198707282011007

Ketua Penguji,

Hatmian Sitohang, M.Kom
NIK. 198503282008002

Tembusan:

1. Arsip Prodi Teknik Informatika
2. Mahasiswa yang bersangkutan