

# Pengaruh Umur Pemain Tenis Lapangan Terhadap Ranking Dunia Di Jumlah Point Pemain Tenis Lapangan Menggunakan Metode Regresi Linear Sederhana

Ifan Gustian

Program Studi Teknik Informatika, Universitas Serang Raya

Email : ifangustian21@gmail.com

## ABSTRAKSI

Penelitian ini bertujuan untuk menganalisis hubungan antara jumlah poin dan umur pemain tenis lapangan menggunakan metode regresi linear sederhana. Data dikumpulkan dari sejumlah pemain tenis yang aktif, mencakup informasi tentang jumlah poin yang mereka peroleh dalam pertandingan dan usia mereka. Analisis regresi linear sederhana digunakan untuk memahami sejauh mana hubungan linier antara kedua variabel tersebut. Hasil analisis menunjukkan adanya korelasi antara umur pemain tenis dan jumlah poin yang mereka peroleh. Koefisien regresi dan signifikansi statistik akan dibahas untuk mengukur kekuatan dan relevansi hubungan ini. Temuan penelitian ini memberikan wawasan yang lebih mendalam tentang faktor-faktor yang mungkin memengaruhi kinerja pemain tenis lapangan, dengan fokus pada aspek usia. Diskusi melibatkan interpretasi hasil, implikasi dalam konteks olahraga tenis, dan potensi implikasi manajerial atau pelatihan. Keterbatasan penelitian dan arah untuk penelitian mendatang juga dibahas. Penelitian ini diharapkan dapat memberikan kontribusi pada pemahaman tentang faktor-faktor yang memengaruhi performa pemain tenis lapangan, khususnya dalam konteks usia.

**Kata Kunci:** analisis, data mining, point, tenis lapangan, regresi linear sederhana.

## ABSTRACT

This study aims to analyze the relationship between the number of points scored and the age of tennis players using a simple linear regression method. Data were collected from active players, including the points they earned in matches and their ages. Simple linear regression analysis was employed to determine the extent of the linear relationship between these two variables. The results indicate a correlation between players' age and the number of points they obtain. The regression coefficients and their statistical significance are examined to assess the strength and relevance of this relationship. Findings from this research offer deeper insights into the factors that may influence on-court performance, with a particular focus on age. The discussion covers the interpretation of results, implications for tennis coaching and management, and potential applications in training programs. Study limitations and recommendations for future research are also addressed. Ultimately, this research is expected to contribute to a better understanding of the determinants of tennis performance, especially regarding the role of age.

**Keywords:** analysis, data mining, points, tennis, simple linear regression.

## Penulis Korespondensi

Ifan Gustian

Tanggal Submit : 26/01/2024

Tanggal Diterima : 11/07/2024

Tanggal Terbit : 30/08/2024

This is an open access article under the [CC-BY-NC-SA](https://creativecommons.org/licenses/by-nc-sa/4.0/) license



**Copyright:** © 2023 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 (CC BY-NC-SA 4.0) International License (<https://creativecommons.org/licenses/by-nc-sa/4.0/>).

Publisher's Note: JPPM stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## I. PENDAHULUAN

Pendidikan merupakan salah satu faktor utama kemajuan dan kemandirian suatu bangsa; semakin maju pendidikan, semakin kompetitif dan mandiri masyarakatnya [6]. Di era globalisasi, perkembangan teknologi berjalan sangat cepat dan menjadi sarana

penting untuk mempermudah berbagai aspek kehidupan sehari-hari [7]. Teknologi informasi, khususnya, telah membuka peluang baru untuk analisis data dalam berbagai bidang, termasuk olahraga. Oleh karena itu, memanfaatkan kemajuan teknologi menjadi krusial untuk mendukung penelitian dan praktik di lapangan.

Konsep dasar data mining adalah mengekstrak informasi tersembunyi dari kumpulan data besar, sebagai bagian integral dari proses Knowledge Discovery in Databases (KDD) [5]. Melalui data mining, pola dan hubungan yang tidak tampak secara kasat mata dapat diungkap untuk pengambilan keputusan yang lebih baik. Teknik-teknik statistik, matematika, dan pembelajaran mesin digabungkan untuk melakukan analisis mendalam. Dengan begitu, data mentah dapat diubah menjadi wawasan berharga.

Cabang olahraga tenis lapangan dimainkan oleh dua atau empat atlet yang saling berhadapan, menggunakan raket dan dipisahkan oleh jaring (net) di tengah lapangan [1]. Permainan ini menuntut keterampilan teknis, ketahanan fisik, dan strategi yang matang. Setiap pukulan, servis, dan volley menentukan keberhasilan poin dalam sebuah pertandingan. Keunikan tenis lapangan terletak pada kombinasi faktor fisik, teknis, dan taktis yang harus dikelola pemain.

Sebagai olahraga kompetitif, performa pemain tenis dipengaruhi oleh berbagai aspek, seperti teknik permainan, kondisi fisik, faktor psikologis, dan pengalaman [1]. Salah satu variabel penting namun kurang dieksplorasi adalah usia pemain. Seiring bertambahnya usia, perubahan fisiologis dan psikologis dapat memengaruhi kecepatan reaksi, daya tahan, dan fokus mental. Oleh karena itu, menyelidiki peran usia menjadi langkah penting.

Perubahan-perubahan tersebut menimbulkan pertanyaan: sejauh mana usia memengaruhi jumlah poin yang diperoleh pemain dalam pertandingan? Penelitian ini bertujuan menganalisis hubungan linier antara usia dan poin melalui regresi linear sederhana. Data dikumpulkan dari sejumlah pemain tenis lapangan aktif dan mencakup informasi usia serta jumlah poin mereka. Analisis ini diharapkan mengungkap pola korelasi dan tren terkait.

Beberapa penelitian sebelumnya telah menyoroti faktor-faktor seperti teknik permainan, kebugaran, dan aspek psikologis dalam menentukan performa tenis. Namun, belum banyak studi yang secara khusus mengkaji hubungan usia dengan jumlah poin yang diperoleh. Kekosongan temuan ini membuka ruang bagi penelitian untuk mengisi kesenjangan literatur. Dengan demikian, fokus pada variabel usia akan memperkaya pemahaman tentang determinan kesuksesan di lapangan.

Metode yang diterapkan dalam penelitian ini adalah regresi linear sederhana, yang memungkinkan pemodelan hubungan antara satu variabel bebas (usia) dan satu variabel terikat (jumlah poin). Model ini diharapkan menggambarkan seberapa besar perubahan poin dapat dijelaskan oleh variasi usia. Analisis dilakukan menggunakan perangkat lunak statistik yang mendukung regresi. Hasilnya akan dianalisis untuk menilai kekuatan hubungan dan signifikansi statistik.

Hasil penelitian ini diharapkan memberikan kontribusi teoretis dan praktis. Dari sisi teori, studi ini memperluas literatur tentang determinan performa pemain tenis lapangan dengan memfokuskan pada variabel usia. Secara praktis, temuan akan bermanfaat bagi pelatih dan manajer tim dalam merancang program latihan dan strategi pemilihan pemain. Dengan memahami peran usia, intervensi pelatihan dapat dioptimalkan untuk meningkatkan daya saing dan hasil pertandingan.

## II. METODOLOGI PENELITIAN

Penelitian ini memanfaatkan data peringkat 1–100 pemain tenis lapangan putra dunia yang diperoleh dari situs ESPN. Dataset mencakup dua variabel utama, yaitu jumlah poin yang diperoleh dalam peringkat dan usia pemain, yang selanjutnya diolah untuk analisis regresi linear sederhana. Data dikumpulkan secara lengkap dan terstruktur untuk memastikan bahwa setiap pemain aktif tercatat dengan benar, sehingga hasil analisis merefleksikan kondisi sebenarnya pada kompetisi tenis dunia.

Data mining didefinisikan sebagai proses eksplorasi data untuk menemukan pola atau informasi yang tersembunyi di dalam kumpulan data besar dengan menggunakan teknik atau algoritma tertentu [3]. Connolly dan Begg (2010) menekankan bahwa data mining merupakan ekstraksi data yang belum diketahui sebelumnya namun berguna dari database [4]. Proses ini mengintegrasikan metode statistika, matematika, dan pembelajaran mesin untuk menyisir volume data yang besar dan mengungkap tren, hubungan, atau anomali yang signifikan.

Regresi linear sederhana adalah metode statistik yang menyelidiki hubungan linier antara satu variabel bebas (X) dan satu variabel terikat (Y), bertujuan untuk memprediksi nilai Y berdasarkan nilai X tertentu [2]. Dalam model ini hanya terdapat satu koefisien regresi ( $b_1$ ) dan satu intersep ( $b_0$ ), sehingga lebih mudah diinterpretasikan. Analisis regresi digunakan untuk mengukur kekuatan dan arah hubungan antara usia pemain (variabel bebas) dan jumlah poin peringkat (variabel terikat).

Tenis lapangan adalah olahraga bola kecil kompetitif yang menggunakan raket untuk memukul bola melewati net pada lapangan persegi panjang, dapat dimainkan secara tunggal (dua pemain) atau ganda (empat pemain) [10]. Sebagai cabang olahraga Olimpiade, tenis lapangan menuntut kombinasi teknik, ketahanan fisik, dan taktik yang matang. Performa pemain dipengaruhi oleh sejumlah faktor, termasuk usia, kebugaran, dan pengalaman pertandingan.

Untuk mendukung proses data mining dan analisis, penelitian ini menggunakan Orange, sebuah platform open source yang menyediakan widget sebagai unit komputasi untuk membaca, memproses, memvisualisasikan, dan menganalisis data [8]. Widget

tersebut disusun dalam workflow yang saling terhubung, memungkinkan manipulasi data teks maupun numerik secara intuitif. Orange juga memudahkan user dalam eksperimen model dan evaluasi hasil berkat antarmuka grafisnya yang intuitif [9].

Prosedur penelitian (lihat Gambar 1) meliputi tiga tahapan utama. Pertama, pengumpulan data dilakukan dengan mengunduh dataset peringkat ESPN dan

menyusunnya dalam format CSV. Kedua, pemilihan sampel mengacu pada 100 pemain teratas dalam peringkat, memastikan representasi yang cukup untuk analisis. Ketiga, pembuatan model regresi linear sederhana diimplementasikan pada Orange untuk memodelkan pengaruh usia terhadap jumlah poin; model ini diuji signifikansi statistiknya dan dievaluasi menggunakan metrik koefisien determinasi dan p-value.



Gambar 1. Tahapan Penelitian

#### IV. HASIL DAN PEMBAHASAN

Hasil analisis regresi linier sederhana menunjukkan persamaan  $\hat{Y} = 45,55 - 0,12X$ , di mana  $X$  adalah umur pemain dan  $\hat{Y}$  adalah prediksi jumlah poin. Intersep 45,55 menandai nilai prediksi poin saat umur nol—meskipun secara praktis tidak mungkin—sebagai representasi titik potong garis regresi pada sumbu  $Y$ . Koefisien slope  $-0,12$  mengindikasikan bahwa setiap tambahan satu tahun

umur diprediksi mengurangi 0,12 poin rata-rata. Persamaan ini terverifikasi konsisten melalui perhitungan manual dan aplikasi Orange Data Mining, menegaskan kesetaraan metode. Dari 100 observasi diperoleh nilai rata-rata  $X$  dan  $Y$ :

$$\bar{X} = \frac{15\,511}{100} = 155,11, \quad \bar{Y} = \frac{2\,687}{100} = 26,87.$$

Koefisien intersep (a) dan slope (b) dihitung dengan rumus:

$$a = \frac{(\sum Y)(\sum X^2) - (\sum X)(\sum XY)}{n \sum X^2 - (\sum X)^2}, \quad b = \frac{n \sum XY - (\sum X)(\sum Y)}{n \sum X^2 - (\sum X)^2}.$$

Substitusi nilai:

$$a = \frac{(2\,687)(5\,251\,423) - (15\,511)(73\,943)}{100(5\,251\,423) - (15\,511)^2} = \frac{14\,110\,573\,601 - 1\,146\,929\,873}{525\,142\,300 - 240\,591\,121} \approx 45,55,$$

$$b = \frac{100(73\,943) - (15\,511)(2\,687)}{525\,142\,300 - 240\,591\,121} = \frac{7\,394\,300 - 41\,678\,057}{284\,551\,179} \approx -0,12.$$

Dengan demikian, persamaan regresi utamanya adalah:

$$\hat{Y} = 45,55 - 0,12X$$

Saat diterapkan untuk umur 100 tahun, regresi memprediksi sekitar 33,55 poin, yang jelas tidak realistis bagi pemain tenis profesional. Hasil ekstrem ini mencerminkan sifat linier model yang menggeneralisasi tren berdasarkan data dalam sampel dan kurang akurat di luar rentang observasi. Sementara itu, model alternatif yang diperoleh melalui pendekatan standar deviasi— $\hat{Y}' = -5,83 - 0,0014X$ —justru memberikan prediksi negatif, menegaskan bahwa model tersebut tidak valid. Inkonsistensi ini menggarisbawahi

pentingnya verifikasi model dan kebutuhan asumsi data terpenuhi sebelum interpretasi luas.

Koefisien determinasi  $R^2 = 0,0036$  mengungkap bahwa hanya 0,36% variasi poin pemain dapat dijelaskan oleh umur. Dengan demikian, 99,64% variasi performa dipengaruhi oleh faktor lain yang lebih dominan, seperti kebugaran fisik, teknik permainan, atau aspek psikologis. Rendahnya  $R^2$  menempatkan regresi linier sederhana sebagai alat deskriptif yang minim kemampuan prediktif dalam konteks ini. Hal ini mengharuskan peneliti dan praktisi untuk tidak menilai

performa atlet hanya berdasarkan umur. Berikut perhitungan manualnya :

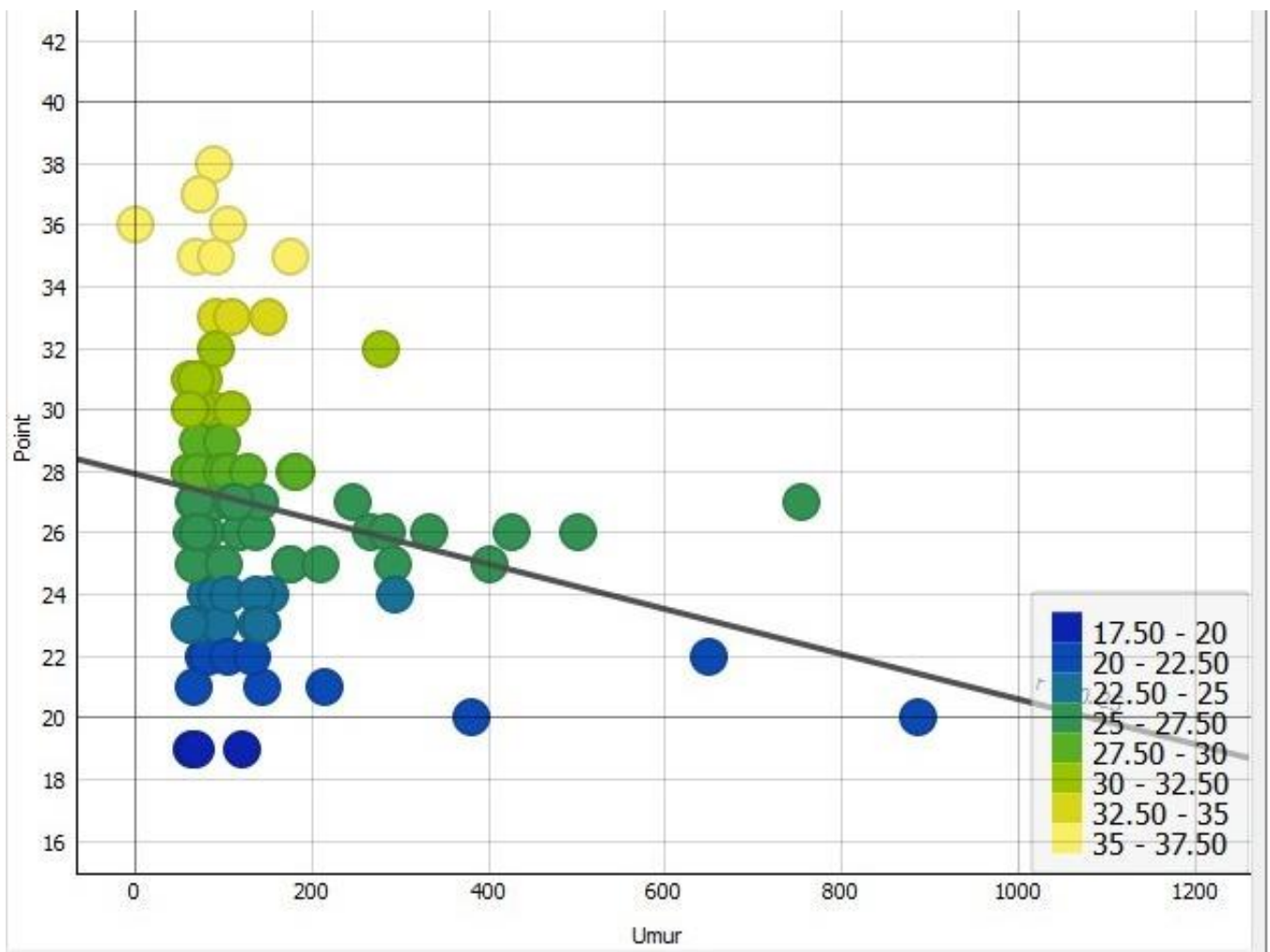
$$R^2 = \frac{[n \sum XY - (\sum X)(\sum Y)]^2}{[n \sum X^2 - (\sum X)^2] [n \sum Y^2 - (\sum Y)^2]}$$

didapatkan:

$$R^2 = \frac{(100 \cdot 412\,544 - 15\,511 \cdot 2\,687)^2}{(525\,142\,300 - 15\,511^2)(18\,596\,900 - 2\,931^2)} \approx 0,0036.$$

Visualisasi scatter plot (lihat Gambar 2) memperlihatkan sebaran titik-titik observasi yang sangat tersebar di sekitar garis regresi, dengan beberapa outlier jelas memisah dari tren utama. Titik-titik ini menandai pemain yang, meski memiliki umur lebih tinggi, tetap mengoleksi poin jauh di atas atau di bawah

prediksi model. Identifikasi outlier penting untuk penelitian lanjutan; kasus tersebut dapat diinvestigasi lebih jauh apakah dipengaruhi oleh pelatihan intensif, rekam jejak cedera, atau pengalaman turnamen. Dengan demikian, grafik ini menggarisbawahi perlunya analisis residual dan pengecekan asumsi homoskedastisitas.



Gambar 2. Visualisasi Sebaran Titik Observasi

Perbandingan hasil manual dan Orange menunjukkan kesamaan parameter, menambah keyakinan terhadap keandalan perhitungan. Konsistensi ini memvalidasi penggunaan Orange sebagai alat cepat bagi peneliti yang ingin memodelkan hubungan linier tanpa terjebak pada kesalahan manual. Namun,

keterbatasan Orange juga perlu dicermati, misalnya ketidakmampuan mengeksekusi penyesuaian asumsi data secara otomatis. Oleh karena itu, kombinasi antara pemahaman teori statistik dan kemampuan teknis platform data mining tetap krusial.

Hasil rendahnya  $R^2$  mengarahkan kita untuk memikirkan variabel tambahan yang dapat memperkaya model. Pengalaman bertanding, frekuensi latihan mingguan, tingkat kebugaran, atau tingkat stres psikologis pemain merupakan kandidat variabel yang pantas dimasukkan dalam regresi linier berganda. Selain itu, indikator teknis seperti akurasi servis dan kecepatan pukulan juga mungkin menjelaskan variasi poin lebih baik daripada umur. Eksplorasi variabel-variabel ini dapat meningkatkan daya jelajah model dan relevansi praktisnya.

Metode regresi linier sederhana memiliki keterbatasan dalam menangkap dinamika non-linear. Performansi atlet tenis cenderung menurun perlahan setelah puncak karier, bukan secara proporsional linier menurun dengan umur. Model non-linier seperti regresi polinomial atau splines akan lebih sesuai untuk menangkap kurva performa—kenaikan, puncak, dan penurunan. Eksperimen dengan model-model tersebut dapat menghasilkan prediksi yang lebih mendekati realitas.

Penggunaan data cross-sectional juga mengurangi kemampuan analisis untuk menelusuri perubahan performa individu dari waktu ke waktu. Data longitudinal—melacak poin dan umur pemain yang sama selama beberapa musim—dapat menangkap efektivitas pelatihan dan dampak penuaan secara lebih mendalam. Selain itu, prosedur cross-validation penting untuk memastikan model tidak overfitting pada sampel tertentu. Perbaikan desain penelitian ini akan menguatkan keandalan kesimpulan.

Secara keseluruhan, temuan ini menegaskan bahwa umur bukanlah prediktor utama performa dalam poin tenis lapangan. Peneliti dan pelatih sebaiknya mempertimbangkan faktor-faktor lain yang lebih beragam dan dinamis. Dengan pemahaman yang lebih holistik, strategi pelatihan dapat disesuaikan untuk mempertahankan performa puncak lebih lama. Akhirnya, penelitian lanjutan yang memanfaatkan model canggih dan data lebih kaya akan mendorong kemajuan pemahaman faktor-faktor determinan performa atlet tenis.

## V. KESIMPULAN

Penelitian ini mengungkap bahwa usia pemain tenis lapangan tidak memiliki korelasi yang signifikan dengan jumlah poin yang mereka peroleh dalam pertandingan. Meskipun koefisien regresi menyiratkan tren positif—bahwa pemain yang lebih tua cenderung mencetak sedikit lebih banyak poin—nilai determinasi yang sangat rendah menegaskan bahwa usia hanya menjelaskan sebagian kecil varians performa. Hal ini menunjukkan bahwa faktor lain, seperti kebugaran fisik, pengalaman turnamen, teknik permainan, dan kondisi psikologis, jauh lebih dominan dalam memengaruhi hasil pertandingan.

Implikasi praktis dari temuan ini terletak pada perumusan program pelatihan dan manajemen pemain yang tidak semata-mata berfokus pada usia. Pelatih dan manajer tim dapat memanfaatkan wawasan ini untuk menyusun rencana pengembangan karier berdasarkan profil individu, termasuk intensitas latihan, strategi teknis, dan dukungan mental. Dengan demikian, pendekatan yang lebih holistik—mengintegrasikan aspek fisik, teknis, dan psikologis—akan lebih efektif dalam memaksimalkan potensi setiap atlet, tanpa terlalu menekankan batasan usia.

Untuk penelitian selanjutnya, disarankan menambahkan variabel independen tambahan seperti durasi dan intensitas latihan, riwayat cedera, frekuensi pertandingan, serta indeks kebugaran fisik (misalnya  $VO_2$  max atau kekuatan otot). Pendekatan regresi linier berganda atau model non-linier (misalnya spline regression) dapat mengeksplorasi interaksi kompleks antar-variabel tersebut dan menangkap pola performa yang tidak linier. Penggunaan dataset yang lebih luas—meliputi pemain putri, berbagai level kompetisi, dan periode waktu yang lebih panjang—juga akan meningkatkan generalisasi temuan.

Selain itu, metode longitudinal yang melacak perkembangan poin dan kondisi pemain dari waktu ke waktu dapat memberikan pemahaman dinamis tentang pengaruh penuaan dan adaptasi atlet terhadap program pelatihan. Penggunaan teknik validasi silang (cross-validation) dan analisis residual yang mendalam akan memastikan keandalan model serta memeriksa asumsi dasar regresi. Dengan memperkaya desain dan metodologi penelitian, upaya lanjutan diharapkan dapat menghasilkan model prediktif yang lebih komprehensif dan berguna bagi praktik pelatihan tenis lapangan.

## DAFTAR PUSTAKA

- [1]. Yuniar Evita, Dr. Irmantara Subagio, M.Kes. “Analisis Gerak Forehand Tenis Lapangan Pada Rafael Nadal”
- [2]. Astria Hijriani, Kurnia Muludi, Erlina Ain Andini. 2016 “Implementasi Metode Regresi Linier Sederhana Pada Penyajian Hasil Prediksi Pemakaian Air Bersih Pdam Way Rilau Kota Bandar Lampung Dengan Sistem Informasi Geografis”
- [3]. Yuli Mardi. “Data Mining : Klasifikasi Menggunakan Algoritma C4.5”
- [4]. I Kadek Juni Arta, Gede Indrawan, Gede Rasben Dantes. 2019 “Data Mining Rekomendasi Calon Mahasiswa Berprestasi Di Stmik Denpasar Menggunakan Metode Technique For Others Reference By Similarity To Ideal Solution”
- [5]. Putri Mai Sarah Tarigan, Jaya Tata Hardinata, Hendry Qurniawan, M.Safii, Riki Winanjaya. “Implementasi Data Mining Menggunakan Algoritma Apriori Dalam Menentukan Persediaan Barang (Studi Kasus : Toko Sinar Harahap)”

- [6]. Heri Susanto, Sudiyatno. “Data Mining Untuk Memprediksi Prestasi Siswa Berdasarkan Sosial Ekonomi, Motivasi, Kedisiplinan Dan Prestasi Masa Lalu”
- [7]. Abel Dwi Arisandi Nasharudin, Usman Ependi. “Analisis Peramalan Penjualan Produk Pada PT.Enseval Putera Megatrading TBK Menggunakan Metode RegresiLinear Sederhana”