

## ANALISA KINERJA MODEL REGRESI DALAM MACHINE LEARNING UNTUK MEMPREDIKSI HARGA BERAS

Muhammad Ardiansyah Sembiring<sup>1</sup>, Febby Wulandari Sembiring<sup>2</sup>  
<sup>1,2</sup>Sistem Informasi, STMIK Royal Kisaran  
e-mail : [adinmantap88@gmail.com](mailto:adinmantap88@gmail.com)<sup>1</sup>, [febywulandari@gmail.com](mailto:febywulandari@gmail.com)<sup>2</sup>

### Abstract

*The trading sector is currently experiencing very significant price increases or decreases. This causes rice to become one of the factors supporting community survival. Based on detailed data from the Indonesian Central Statistics Agency, wholesale rice prices in November 2022 jumped to IDR 11,012/kg, a significant increase of 0.6% compared to the previous month's IDR 10,947/kg. An increase in the price of rice can result in a decrease in people's purchasing power for other needs. So, to anticipate an increase in rice prices, predictions are made using machine learning using a comparison of 7 methods. The purpose of this research is to analyze the performance of regression models in machine learning for predicting rice prices. Ultimately, we will identify the method that yields the most accurate predictions. The estimation employs the regression method, which encompasses seven specific techniques: Linear Regression, Linear Support Vector Regression, RBF Support Vector Regression, Decision Tree Regression, Random Forest Regressor, Gradient Boosting Regression, and NLP Regressor. The best model from the results of this research analysis is the Decision Tree Regression method with model accuracy testing, namely at a data testing ratio of 80:20 of 100%, 100% at a data testing ratio of 70:30 and at a ratio of 60:40 is 100%.*

**Keywords:** Rice Prices, Machine Learning, Predictions

### Abstrak

Sektor perdagangan saat ini mengalami kenaikan atau penurunan harga yang sangat signifikan. Hal ini menyebabkan beras menjadi salah satu faktor penunjang keberlangsungan hidup masyarakat. Berdasarkan data Badan Pusat Statistik Indonesia secara rinci, Harga grosir beras pada November 2022 melonjak hingga Rp 11.012/kg, naik signifikan 0,6% dibandingkan bulan sebelumnya Rp 10.947/kg. Kenaikan harga beras dapat mengakibatkan penurunan daya beli masyarakat terhadap kebutuhan lainnya. Sehingga untuk mengantisipasi adanya kenaikan harga beras dilakukannya prediksi menggunakan machine learning dengan menggunakan perbandingan 7 metode. Tujuan penelitian ini adalah untuk menganalisa kinerja model-model regresi terbaik dalam machine learning yang dapat digunakan untuk memprediksi harga beras. Akan diidentifikasi suatu metode yang menghasilkan prediksi harga beras yang paling akurat. Estimasi ini menggunakan metode regresi yang mencakup tujuh teknik tambahan: Regresi Linier, Regresi Vektor Pendukung Linier, Regresi Vektor Pendukung RBF, Regresi Pohon Keputusan, Regresi Hutan Acak, Regresi Peningkat Gradien, Regresor NLP. Berdasarkan analisis penelitian ini, model terbaik adalah metode Regresi Pohon Keputusan, yang mencapai akurasi model luar biasa sebesar 100% pada berbagai rasio pengujian data: 80:20, 70:30, dan 60:40.

**Kata Kunci:** Harga Beras, Machine Learning, Prediksi

## 1. PENDAHULUAN

Di era globalisasi saat ini, pertumbuhan teknologi informasi yang eksponensial mengubah tatanan kehidupan manusia. Pengaruhnya yang luas meresap ke dalam setiap aspek kehidupan, menawarkan solusi yang belum pernah ada sebelumnya bahkan untuk tantangan yang paling rumit sekalipun. Tidak terkecuali di dalam aspek perdagangan. Sektor perdagangan saat ini sangat sulit ditebak, terkadang mengalami kenaikan atau penurunan harga yang sangat signifikan. Sehingga hal ini menyebabkan bertambahnya kejenuhan di masyarakat. Apa lagi jika kenaikan tersebut menyangkut bahan pokok yang sangat dibutuhkan oleh masyarakat. Salah satu contoh bahan pokok yang paling dibutuhkan masyarakat Indonesia adalah beras.

Beras merupakan salah satu bahan pokok di setiap negara. Hal ini menyebabkan beras menjadi salah satu faktor penunjang keberlangsungan hidup masyarakat. Kenaikan harga beras dapat mengakibatkan penurunan daya beli masyarakat terhadap kebutuhan lainnya (Beli *et al.*, 2024). Berdasarkan data Badan Pusat Statistik Indonesia secara rinci, Harga grosir beras pada November 2022 melonjak hingga Rp 11.012/kg, naik signifikan 0,6% dibandingkan bulan sebelumnya Rp 10.947/kg. Dibandingkan periode yang sama tahun lalu, lonjakannya mencapai 6,14%. Kebanyakan metode atau cara yang digunakan instansi, grosir dan para pedagang kecil lainnya masih memprediksi harga beras dengan menggunakan cara-cara yang tergolong sederhana yaitu hanya berdasarkan pengalaman, selain itu juga berdasarkan prediksi dari beberapa bulan sebelumnya. Metodologi ini mengabaikan berbagai faktor lain yang berpotensi mempengaruhi fluktuasi harga beras, sehingga menjadikannya tidak efektif dalam memprediksi harga secara akurat.

Metode yang sama juga masih digunakan di pengambil keputusan di Kabupaten Asahan. Sehingga menyebabkan tingkat efisiensi hasil prediksi harga beras yang mereka lakukan sangatlah minim dan juga sulit menerka harga beras jika terjadi situasi dan kondisi tertentu yang mungkin dapat mempengaruhi atau bahkan sangat mempengaruhi kenaikan atau penurunan harga beras tersebut. Salah satu tugas utama Perum BULOG adalah menjaga kestabilan harga pangan di pasar dengan cara melakukan intervensi pasokan dan permintaan. Mengingat besarnya tanggung jawab ini, sangat penting bagi Perum BULOG untuk mengadopsi metodologi yang melampaui pendekatan-pendekatan sebelumnya dalam hal efektivitas dan efisiensi. Salah satu metode yang cocok untuk diintegrasikan ke dalam sistem informasi mereka adalah analisis regresi.

Analisis regresi berfungsi sebagai alat statistik untuk meneliti korelasi antara berbagai variabel (Mulyana and Marjuki, 2022). Ini mengkuantifikasi hubungan-hubungan ini secara matematis, enggambarkannya dalam bentuk fungsional. Pada dasarnya, regresi merupakan landasan statistik yang bertujuan untuk membedah dinamika sebab-akibat antara berbagai variabel (Muhammad Bangkit Riksa Utama and Hajarisman, 2021). Kegunaannya mencakup untuk menggambarkan sifat dan kekuatan hubungan antar variabel, sekaligus memfasilitasi prediksi variabel yang belum dipetakan berdasarkan observasi historis dan faktor lainnya. Melalui pembuatan sistem atau aplikasi berbasis web untuk prediksi kenaikan harga beras, maka dapat mempermudah pihak BULOG untuk memprediksi harga beras. Penelitian terkait prediksi harga beras dilakukan dengan menggunakan regresi linier (Arinal and Azhari, 2023). Akan tetapi model yang digunakan belum menggunakan beberapa persentasi data training dan testing yaitu hanya dengan rasio 70% training dan 30% testing. Sedangkan dalam penelitian ini, akan dilakukan menggunakan beberapa metode dan rasio training dan teting dengan tujuan mendapatkan model dengan performance yang lebih baik.

## 2. METODE PENELITIAN

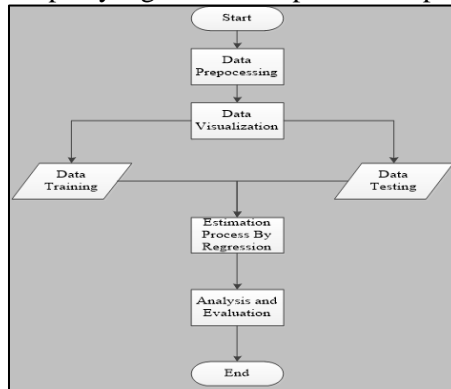
### a. Variabel Riset

Dalam ranah pembelajaran mesin, penelitian ini termasuk dalam kategori supervised learning, sebuah pendekatan pemodelan di mana data yang dipergunakan telah dilengkapi dengan label, target, atau kelas. Sasarannya adalah untuk mengidentifikasi hubungan sebab-akibat antara variabel independen dan variabel target atau berlabel. Variabel yang dipakai dalam studi ini

meliputi Harga Gabah Kering (Rp/Kg), Produksi Padi (Ton), Harga BBM (Liter/Rp), dan Luas Lahan (Hektar), dengan variabel tergantungnya adalah Harga Beras (Rp/Kg).

b. Tahapan–Tahapan Riset

Untuk memvisualisasikan langkah-langkah penelitian ini secara lebih terperinci, akan dibuat sebuah flowchart yang menggambarkan proses yang akan dilalui. Hal ini akan memudahkan pemahaman terhadap setiap tahapan yang akan ditempuh dalam penelitian ini.



Gambar 1. Flowchart Sistem Estimasi

c. Dataset

Data yang akan dipergunakan merupakan data harga beras yang diperoleh dari Badan Pusat Statistik (BPS), meliputi rentang tahun 2019 hingga 2022 dengan jumlah total 97 baris. Terdapat lima kolom dalam dataset ini, mencakup harga gabah kering, luas lahan sawah, harga BBM, produksi per ton, dan harga minyak goreng, yang akan dijadikan sebagai variabel uji coba. Selain itu, terdapat satu kolom yang berperan sebagai target atau label, yakni Harga Beras di Grosir. Untuk tujuan ilustrasi dalam penelitian ini, sampel dataset akan direpresentasikan dengan lima baris teratas seperti yang tampak dalam Tabel(1).

Tabel 1. Sampel Dataset

Harga Gabah Kering	Luas Lahan Sawah	Harga BBM	Produksi Perton	Harga Minyak goreng	Harga Beras Di Grosir
9433.17	4494.67	433043	8500	2514020	13700
9531.20	4501.84	433043	8500	2519020	13700
9595.89	4210.54	433043	8500	2529020	13700
9424.65	4010.54	433043	8500	2517020	13700
9413.91	4209.36	433043	8500	2519020	13700

d. Split Data

Dalam riset ini, pembagian data dilakukan dengan memisahkan data training dan data testing dalam perbandingan 90:10. Data training mencakup 90% dari keseluruhan dataset, sementara data testing mencakup 10%. Algoritma yang menggunakan perbandingan ini menunjukkan performa yang unggul dibandingkan dengan algoritma lainnya, dengan tingkat akurasi mencapai 99,40% [2]. Selain itu, percobaan juga dilakukan dengan rasio pembagian data 80:20, 70:30, dan 60:40 untuk menemukan model regresi dengan hasil estimasi optimal.

e. Implementasi Model Regresi

Penggunaan regresi merupakan pendekatan yang umum dalam menganalisis hubungan antara variabel x (variabel independen) dan variabel y (variabel dependen). Dalam proses estimasi

ini, digunakan tujuh model regresi yang berbeda dengan maksud untuk membandingkan dan menentukan model yang paling optimal untuk digunakan.

#### 1. Model Regesi Linier

Regresi linear sederhana merupakan model yang memeriksa korelasi antara satu variabel prediktor dan satu variabel respons (Hasna and Achmad, 2022). Persamaan yang digunakan dalam regresi linear sederhana adalah sebagai berikut:

$$Y = a + bx$$

Dalam konteks ini, Y mewakili variabel dependen, sedangkan a merupakan titik potong garis pada sumbu Y, dan b adalah koefisien variabel independen (x). Variabel x sendiri adalah variabel independen dalam model.

Menurut sumber (Ronaldi, 2022) regresi linear berganda, atau yang dikenal sebagai multiple linear regression, adalah suatu model analisis yang menggambarkan hubungan antara variabel respons (Y) dengan beberapa variabel independen (x), yang jumlahnya lebih dari satu. Persamaan yang digunakan dalam regresi berganda adalah:

$$Y = b_0 + b_1X_1 + b_2X_2 + b_3X_3 + \dots + b_nX_n$$

Dalam penjelasan ini, Y merujuk pada variabel respon, sedangkan X adalah variabel prediktor yang mempengaruhi variabel respon.  $b_0$  adalah intercept, dan b adalah koefisien dalam hubungan tersebut.

#### 2. Model Support Vector Regression – Linear

Sebagai salah satu model regresi, support vector regression dipilih untuk mengatasi masalah overfitting dan mencapai akurasi yang baik (Oktafiani and Rianto, 2023). Overfitting terjadi ketika model terlalu diperhatikan terhadap data pelatihan (A. Rahim, Ingrid Yanuar Risca Pratiwi and Muhammad Ainul Fikri, 2023), sehingga prediksinya cenderung menjadi hampir sempurna. Persamaan yang digunakan dalam support vector regression adalah sebagai berikut:

$$f(x) = w^t(p(x)) + b$$

Dalam penjelasan ini,  $f(x)$  menggambarkan fungsi regresi, sedangkan  $w^t$  merupakan vektor bobot dengan dimensi 1.  $p(x)$  menandakan titik dalam ruang F, hasil dari pemetaan x pada ruang input. Sementara itu, b adalah bias yang memengaruhi prediksi.

#### 3. Support Vector Regression dengan kernel Gaussian-RBF

Dalam metode Support Vector Regression dengan kernel Gaussian-RBF (Suryani and Mustakim, 2022), persamaan yang digunakan adalah::

$$K(x_i, x) = \exp\left(-\frac{1}{2a^2} \|x - x_i^T\|^2\right) \quad (2)$$

#### 4. Decision Tree Regresi

Decision tree regression pertama-tama menghitung nilai entropi, seperti yang dijelaskan dalam persamaan berikut (Anggara, Widjaja and Suteja, 2022):

$$entro(S) = \sum_{i=1}^m -p(w_i|S) \cdot \log_2(w_i|S)$$

Setelah mengidentifikasi variabel seperti S untuk Himpunan Kasus, M untuk Total Kelas Data, dan  $(w_i|S)$  untuk Proporsi kelas ke-i dalam semua data latih yang diproses di node S, langkah selanjutnya adalah mencari nilai gain.

$$Gain(S, J) = Entropy(S) - \sum_{i=j}^n p(v_i|S) * E(S_i)$$

Variabel S merupakan Himpunan Kasus, J untuk Fitur, n untuk Banyak Kelas dalam node (akar),  $(v_i|S)$  untuk Proporsi nilai v yang muncul pada kelas dalam node (akar), dan  $(S_i)$  untuk Entropi komposisi nilai v dari kelas ke-j dalam data node ke-i

### 5. Random Forest Regresi

Random forest merupakan algoritma regresi yang memanfaatkan teknik yang menggabungkan prediksi dari berbagai pemecahan dalam machine learning (Aditya Pratama *et al.*, 2024). Pendekatan ini bertujuan untuk menghasilkan prediksi dengan akurasi yang lebih baik dibandingkan dengan model tunggal. Persamaan random forest regressor dapat dilihat di bawah ini (Firmansyah *et al.*, 2022):

$$\hat{y}_i = \frac{1}{N_{\text{tree}}} \sum_{n=1}^{N_{\text{tree}}} \hat{y}_n$$

Keterangan :

$\hat{y}_i$  = Estimasi

$N_{\text{tree}}$  = Jumlah Tree (pohon yang ditemukan)

$\hat{y}_n$  = Estimasi tree ke n

### 6. Gradient Boosting Regression

Sebagai elemen dalam algoritma ensemble, gradient boosting regression bertujuan untuk meningkatkan akurasi prediksi (Satria, Badri and Safitri, 2023). Model ini memiliki kemampuan untuk menangani pola data yang kompleks dengan memanfaatkan struktur pohon keputusan. Persamaan gradient boosting regression adalah sebagai berikut (Fitri, 2023):

$$r_{im} = - \left[ \frac{\partial L(y_i, F(x_i))}{\partial F(x_i)} \right]_{F(x)=F_{m-1}(x)} \quad \text{for } i=1, \dots, n$$

### 7. Natural Language Processing Regresi

Pengolahan bahasa alami (Natural Language Processing/NLP) merujuk pada kapasitas program komputer dalam memahami dan mengolah bahasa manusia (Azriansyah, Indra and Azriansyah, 2023), baik secara lisan maupun tulisan. Tujuan utamanya adalah untuk menciptakan atau mengembangkan aplikasi yang memfasilitasi interaksi antara manusia dan mesin melalui penggunaan bahasa alami.

#### f. Performane

Setelah data diolah menggunakan model-model diatas, langkah berikutnya adalah mengevaluasi akurasi dengan mempertimbangkan nilai error. Dalam proses estimasi data, di mana ketidakpastian hadir, penting untuk memperhitungkan nilai error sebagai indikator akurasi estimasi. Nilai error diukur sebagai selisih antara nilai observasi yang sebenarnya dengan nilai prediksi. Evaluasi dilakukan dengan metode estimasi seperti MAE, MSE, dan RMSE. Semakin kecil nilainya, semakin baik estimasinya. Sebaliknya, R2-Score yang mendekati 1 menandakan bahwa variabel independen memberikan informasi yang baik untuk memprediksi variabel dependennya (Anggi Mei Sarah1, Bambang Kurniadi, 2023).

## 3. HASIL DAN PEMBAHASAN

Fokus utama penelitian ini adalah untuk memperjelas keterkaitan antara variabel x dan variabel y. Selain itu, data akan diuji melalui empat skenario pembagian data yang berbeda, menjadikannya lebih ekstrem., masing-masing diuji dengan tujuh model regresi yang berbeda dan diukur menggunakan empat metode evaluasi yang berbeda pula. Metode regresi dipilih untuk membandingkan tujuh metode regresi yang berbeda, dengan harapan dapat menemukan model terbaik untuk memprediksi kenaikan harga beras. Penelitian ini akan melakukan pembagian data ke dalam lima skenario, termasuk rasio 90:10, 80:20, 70:30, dan 60:40, untuk menguji konsistensi model regresi dalam menghasilkan estimasi yang akurat. Dengan hasil penelitian ini, diharapkan

dapat ditemukan model regresi yang paling sesuai untuk memprediksi kenaikan harga beras. Untuk pengujian pertama dengan rasio 90:10, di mana 90% training dan dari 10% testing

Tabel 2. Rasio pengujian training testing 90:10

MODEL	MAE	MSE	RMSE	R2-SCORE
LinearRegression	2.57097342	8.09162808	2.84457872	8.76741175
SVR-Linear	5.75362895	5.70303422	7.55184363	-8.68735538
SVR-RBF	6.53672758	6.70439866	8.18803924	-2.12732118
DecisionTree	0.00000000	0.00000000	0.00000000	1.00000000
RandomForest	2.83946667	1.71765903	4.14446501	9.97961519
GradientBoosting	1.95821820	6.09467351	2.46873925	9.99276697
MLP-Regressor	9.25379084	9.22604168	9.60522862	-1.09482694

Dalam mengevaluasi hasil estimasi, penting untuk memperhatikan nilai-nilai yang ekstrem. Kriteria yang digunakan adalah memilih nilai-nilai yang paling ekstrem, baik itu nilai-nilai terkecil atau terbesar, sesuai dengan metrik evaluasi yang digunakan. Dalam kasus ini, kita mencari nilai-nilai terkecil untuk MAE, MSE, dan RMSE, sementara untuk R2-Score, kita mencari nilai yang paling mendekati 1 atau paling jauh dari 0, menandakan hasil estimasi yang paling baik. Dari tabel yang disediakan, metode Decision Tree Regressor menunjukkan konsistensi dalam memberikan hasil estimasi terbaik, sehingga dapat dianggap sebagai metode regresi yang paling ekstrem dalam hal kualitas estimasi, terutama pada rasio data 90:10. Selanjutnya rasio pengujian 80:20, di mana 80% dan testing 20%.

Tabel 3. Rasio pengujian training testing 80:20

MODEL	MAE	MSE	RMSE	R2-SCORE
LinearRegression	2.16944988	1.07379982	3.27688849	8.48640493
SVR-Linear	3.24875394	1.37526335	3.70845432	-1.93851876
SVR-RBF	7.10848208	7.13386877	8.44622328	-5.56812088
DecisionTree	0.00000000	0.00000000	0.00000000	1.00000000
RandomForest	3.07801857	2.92292682	1.75442827	9.97961519
GradientBoosting	1.95821820	6.09467351	1.70965693	9.99587993
MLP-Regressor	9.25379084	9.22604168	9.60522862	-1.09482694

Dalam pengujian kali ini diperoleh metode *Decision Tree Regressor* sebagai metode regresi terbaik diantara 7 metode lainnya.

Tabel 4. Rasio pengujian training testing 70:30

MODEL	MAE	MSE	RMSE	R2-SCORE
LinearRegression	5.32009248	3.97633510	6.30581882	4.64209317
SVR-Linear	5.55608873	4.78066883	6.91423809	-6.44169512
SVR-RBF	6.54412376	7.53998427	8.68330828	-1.59746202
DecisionTree	0.00000000	0.00000000	0.00000000	1.00000000
RandomForest	6.81670000	2.69172031	1.6406431	9.97961519
GradientBoosting	3.02796248	1.38585567	3.7227082	9.98123631

MLP-Regressor	1.80799560	4.09946923	2.02471460	-5.51382373
---------------	------------	------------	------------	-------------

Dalam pengujian dengan rasio 70:30, metode Decision Tree Regressor kembali menonjol sebagai pilihan terbaik di antara tujuh model regresi lainnya. Metode decision tree regressor mendominasi sebagai pilihan terbaik dari tujuh metode regresi lainnya. Terlihat dari nilai estimasi yang paling baik pada ketiga model pengukuran, metode Decision Tree Regressor menonjol sebagai pilihan terbaik dalam pengujian dengan rasio 70:30, dengan hasil estimasi yang mendekati 1. Dengan demikian, metode ini dapat dianggap sebagai yang terbaik dalam pengujian dengan akurasi 70:30.

Tabel 5. Rasio pengujian training testing 60:40

MODEL	MAE	MSE	RMSE	R2-SCORE
LinearRegression	4.77820065	3.09962932	5.56743147	5.53415413
SVR-Linear	4.98589259	5.04809256	7.10499301	2.49553466
SVR-RBF	7.21266244	7.03352728	8.38645174	-1.33290081
DecisionTree	0.00000000	0.00000000	0.00000000	1.00000000
RandomForest	8.81392727	2.98367919	1.72733297	9.99102734
GradientBoosting	1.92599802	6.22769324	2.49553466	1.16351405
MLP-Regressor	1.4900487	2.67368464	1.63514056	-3.84215853

Berdasarkan tabel yang tercantum di atas, dapat dipastikan bahwa metode Decision Tree Regressor merupakan pilihan terbaik dalam pengujian dengan rasio 60:40, sekaligus menjadi metode yang paling dominan dengan nilai estimasi yang optimal. Decision Tree Regressor memperoleh nilai 0 pada pengukuran MAE, 0 pada MSE, 0 pada RMSE, dan 1 pada R2-Score. Dengan demikian, pada pengujian data dengan rasio 60:40, Decision Tree Regressor dapat disimpulkan sebagai metode terbaik.

#### Analisis Perbandingan Metode

Metode terbaik yang paling konsisten dalam menghasilkan nilai estimasi terbaik adalah Decision Tree Regression. Model ini mencapai akurasi yang tinggi pada pengujian data: 100% pada rasio 90:10, 100% pada rasio 80:20, 100% pada rasio 70:30, dan juga 100% pada rasio 60:40. Dengan tingkat akurasi yang ekstrem dan konsisten ini, metode Decision Tree Regression menjadi pilihan yang sangat kuat untuk memprediksi harga beras. Dibandingkan dengan penelitian terdahulu (Arinal and Azhari, 2023), model yang digunakan adalah regresi, sedangkan dalam penelitian ini model yang terbaik yang dihasilkan berasal dari metode Decision Tree.

#### 4. KESIMPULAN

Kesimpulan dari penelitian ini sangat jelas: metode Decision Tree Regression adalah pilihan terbaik untuk memprediksi harga beras. Dengan akurasi yang baik dalam keempat skenario rasio data training dan testing yang berbeda, tidak ada keraguan bahwa Decision Tree Regression menonjol sebagai metode yang paling efektif. Oleh karena itu, implementasi model ini dalam pembuatan dan deploy aplikasi untuk memprediksi harga beras di Kabupaten Asahan merupakan langkah yang sangat tepat. Untuk penelitian selanjutnya, disarankan untuk mengumpulkan dataset yang lebih besar guna mendapatkan wawasan yang lebih komprehensif dan hasil yang lebih kuat untuk masa depan.

## 5. DAFTAR PUSAKA

- A. Rahim, A.M., Inggrid Yanuar Risca Pratiwi and Muhammad Ainul Fikri (2023) “Klasifikasi Penyakit Jantung Menggunakan Metode Synthetic Minority Over-Sampling Technique Dan Random Forest Classifier,” *Indonesian Journal of Computer Science*, 12(5), pp. 2995–3011. Available at: <https://doi.org/10.33022/ijcs.v12i5.3413>.
- Aditya Pratama, M. *et al.* (2024) “Perbandingan Performa Algoritma Linear Regresi dan Random Forest untuk Prediksi Harga Bawang Merah di Kota Samarinda,” *Jurnal Ilmu Teknik*, 1(2), pp. 172–182. Available at: <https://doi.org/10.62017/tekonik>.
- Anggara, E.D., Widjaja, A. and Suteja, B.R. (2022) “Prediksi Kinerja Pegawai sebagai Rekomendasi Kenaikan Golongan dengan Metode Decision Tree dan Regresi Logistik,” *Jurnal Teknik Informatika dan Sistem Informasi*, 8(1), pp. 218–234. Available at: <https://doi.org/10.28932/jutisi.v8i1.4479>.
- Anggi Mei Sarah1, Bambang Kurniadi, E.W. (2023) “IMPLEMENTASI METODE REGRESI LINEAR DALAM MEMREDIKSI PENYAKIT ANEMIA SECARA DINI,” *IMPLEMENTASI METODE REGRESI LINEAR DALAM MEMREDIKSI PENYAKIT ANEMIA SECARA DINI*, 3(1), pp. 14–23.
- Arinal, V. and Azhari, M. (2023) “Penerapan Regresi Linear Untuk Prediksi Harga Beras Di Indonesia,” *Jurnal Sains dan Teknologi*, 5(1), pp. 341–346.
- Azriansyah, N., Indra, E. and Azriansyah, N. (2023) “Penerapan Natural Language Processing Untuk Analisis Sentimen Terhadap Aplikasi Streaming,” *Jurnal Ilmiah BETRIK (Besemah Teknologi Informasi dan Komputer)*, 14(2), pp. 273–282. Available at: <https://ejournal.pppmitpa.or.id/index.php/betrik/article/view/96>.
- Beli, D. *et al.* (2024) “Pengaruh Naiknya Harga Beras Terhadap Pendapatan Riil dan Daya Beli Pangan Di Kelurahan Cilangkap,” *jupensal*, 1(2), pp. 311–318.
- Firmansyah, I. *et al.* (2022) “Komparasi Random Forest Dan Logistic Regression Dalam Klasifikasi Penderita Covid-19 Berdasarkan Gejalanya,” *Journal of Science and Social Research*, 5(3), p. 595. Available at: <https://doi.org/10.54314/jssr.v5i3.994>.
- Fitri, E. (2023) “Analisis Perbandingan Metode Regresi Linier, Random Forest Regression dan Gradient Boosted Trees Regression Method untuk Prediksi Harga Rumah,” *Journal of Applied Computer Science and Technology*, 4(1), pp. 58–64. Available at: <https://doi.org/10.52158/jacost.v4i1.491>.
- Hasna and Achmad, A.I. (2022) “Metode Regresi Probit Biner untuk Pemodelan Faktor-Faktor yang Mempengaruhi Diagnosis Penyakit Jantung,” *Jurnal Riset Statistika*, pp. 28–34. Available at: <https://doi.org/10.29313/jrs.vi.721>.
- Muhammad Bangkit Riksa Utama and Hajarisman, N. (2021) “Metode Pemilihan Variabel pada Model Regresi Poisson Menggunakan Metode Nordberg,” *Jurnal Riset Statistika*, 1(1), pp. 35–42. Available at: <https://doi.org/10.29313/jrs.v1i1.24>.
- Mulyana, D.I. and Marjuki (2022) “Optimasi Prediksi Harga Udang Vaname Dengan Metode Rmse Dan Mae Dalam Algoritma Regresi Linier,” *Jurnal Ilmiah Betrik*, 13(1), pp. 50–58. Available at: <https://doi.org/10.36050/betrik.v13i1.439>.
- Oktafiani, R. and Rianto, R. (2023) “Perbandingan Algoritma Support Vector Machine (SVM) dan Decision Tree untuk Sistem Rekomendasi Tempat Wisata,” *Jurnal Nasional Teknologi dan Sistem Informasi*, 9(2), pp. 113–121. Available at: <https://doi.org/10.25077/teknosi.v9i2.2023.113-121>.
- Ronaldi, R. (2022) “Keputusan Pembelian: Brand Image Dan Brand Awareness Produk Smartphone Samsung,” *JEM: Jurnal Ekonomi dan Manajemen STIE Pertiba Pangkalpinang*, Vol. 9 No.(Juli-Des), pp. 1–10.
- Satria, A., Badri, R.M. and Safitri, I. (2023) “Prediksi Hasil Panen Tanaman Pangan Sumatera dengan Metode Machine Learning,” *Digital Transformation Technology*, 3(2), pp. 389–398. Available at: <https://doi.org/10.47709/digitech.v3i2.2852>.
- Suryani, S. and Mustakim, M. (2022) “Estimasi Keberhasilan Siswa dalam Pemodelan Data



Berbasis Learning Menggunakan Algoritma Support Vector Machine,” *Bulletin of Informatics and Data Science*, 1(2), p. 81. Available at: <https://doi.org/10.61944/bids.v1i2.36>.