

Implementasi Data Mining Untuk Memprediksi Konsumsi Bahan Bakar Menggunakan Metode Regularized Linear Regression

Implementation of Data Mining to Predict Fuel Consumption Using The Regularized Linear Regression Method

Elvies Kamti^a Raynaldi Felim^b Irvin Chandra^c Yennimar^{d*}

Universitas Prima Indonesia Medan^{a,b,c,d}
elvieskamti2310@gmail.com

Abstract

The rapid development of the automotive industry in the era of globalization has increased competition and driven the need for vehicle fuel efficiency. Fuel consumption prediction is crucial to reduce operational costs, reduce carbon emissions, and support sustainable environmental policies. This study aims to build a vehicle fuel consumption prediction model using the Regularized Linear Regression method by utilizing vehicle data such as Vehicle Class, Engine Size, Cylinders, Transmission, and CO2 Emissions. The model was built based on 639 data from 36 car brands, including Suzuki, Honda, Audi, and Toyota. The analysis results show that this method is able to provide an accuracy rate of 79% based on the R-squared value and has a Mean Squared Error (MSE) of 2.01, which indicates quite good prediction performance. This approach is compared with several previous studies and shows improvements in the presentation of accuracy and measurement of prediction errors. This study makes a significant contribution to the development of a more accurate fuel consumption prediction system, which can be used by vehicle manufacturers, transportation companies, and policymakers to improve the efficiency and sustainability of the transportation sector.

Keywords: Fuel Consumption Prediction, Regularized Linear Regression, Data Mining, Energy Efficiency, Carbon Emissions

Abstrak

Pesatnya perkembangan industri otomotif di era globalisasi telah meningkatkan persaingan dan mendorong kebutuhan akan efisiensi bahan bakar kendaraan. Prediksi konsumsi bahan bakar menjadi krusial untuk menekan biaya operasional, mengurangi emisi karbon, serta mendukung kebijakan lingkungan yang berkelanjutan. Penelitian ini bertujuan untuk membangun model prediksi konsumsi bahan bakar kendaraan menggunakan metode Regularized Linear Regression dengan memanfaatkan data kendaraan seperti Vehicle Class, Engine Size, Cylinders, Transmission, dan CO Emissions. Model dibangun berdasarkan 639 data dari 36 merek mobil, termasuk Suzuki, Honda, Audi, dan Toyota. Hasil analisis menunjukkan bahwa metode ini mampu memberikan tingkat akurasi sebesar 79% berdasarkan nilai R-squared dan memiliki Mean Squared Error (MSE) senilai 2.01, yang menunjukkan performa prediksi yang cukup baik. Pendekatan ini dibandingkan dengan beberapa penelitian sebelumnya dan menunjukkan peningkatan dalam penyajian akurasi serta pengukuran kesalahan prediksi. Penelitian ini memberikan kontribusi yang cukup besar dalam perkembangan sistem prediksi konsumsi bahan bakar yang lebih akurat, yang dapat digunakan oleh produsen kendaraan, perusahaan transportasi, dan pembuat kebijakan untuk meningkatkan efisiensi dan keberlanjutan sektor transportasi.

Kata Kunci: Prediksi Konsumsi Bahan Bakar, Regularized Linear Regression, Data Mining, Efisiensi Energi, Emisi Karbon

1. Pendahuluan

Pertumbuhan industri yang signifikan pada era globalisasi saat ini menyebabkan daya saing antar perusahaan di Indonesia menjadi lebih kuat (Nunka & Pranoto, 2024). Data mining merupakan proses pencarian data yang cukup besar dari sekumpulan

data yang besar, sedangkan regresi linier ialah suatu cara statistik yang berfungsi untuk membuat model suatu hubungan linear antar variabel dependen dengan variabel independen baik itu satu maupun banyak, sedangkan k-nearest neighbor merupakan algoritma yang berfungsi untuk melakukan pengelompokan dan regresi (Kurniawan, Rahaningsih, & Suprpti, 2024). Data mining adalah tentang metode untuk penggalian pengetahuan atau menemukan pola dari suatu data (Arinal & Azhari, 2023).

Seiring dengan meningkatnya jumlah kendaraan di dunia, konsumsi bahan bakar ialah salah satu komponen fundamental yang perlu diperhatikan dalam pengelolaan industri otomotif dan transportasi. Kenaikan harga bahan bakar, kebutuhan untuk mengurangi emisi karbon, serta tantangan lingkungan memaksa produsen kendaraan dan perusahaan transportasi untuk mencari solusi yang lebih efisien dalam penggunaan bahan bakar. Prediksi konsumsi bahan bakar yang lebih akurat dapat membantu dalam merancang kebijakan yang lebih efektif dan mengoptimalkan operasi kendaraan untuk mengurangi biaya dan dampak negatif terhadap lingkungan. Namun, dalam praktiknya, prediksi konsumsi bahan bakar sering kali hanya didasarkan pada data historis atau perkiraan kasar, yang tidak selalu mencerminkan kondisi nyata dan variabel dinamis yang mempengaruhi konsumsi. Oleh karena itu, dibutuhkan pendekatan yang lebih canggih untuk memprediksi konsumsi bahan bakar dengan menggunakan metode analisis data yang lebih akurat, seperti data mining. Prediksi merupakan suatu ilmu untuk menduga kejadian yang akan terjadi menggunakan kumpulan data yang diolah sedemikian rupa dan dapat memberikan proyeksi jawaban yang akan datang dengan menggunakan model yang sistematis (Indarwati, Irawati, & Rimawati, 2019).

Metode yang digunakan dalam penelitian ini yaitu metode regresi linier sederhana (Husdi & Dalai, 2023). Metode Regularized Linear Regression dipilih dalam penelitian ini karena mampu mengatasi masalah multikolinearitas dalam data dan mengurangi overfitting, yang merupakan tantangan umum dalam pemodelan dengan data yang kompleks. Regresi adalah salah satu cara untuk membuat suatu model yang dapat digunakan untuk melakukan prediksi suatu nilai dari masukan yang diberikan (Deni & Latifah, 2021). Penerapan data model algoritma regresi linier sederhana dapat terlaksana dengan baik dan dapat memberikan pengetahuan baru (Suwaryo et al., 2023).

Variabel yang digunakan dalam penelitian ini seperti Vehicle Class, Engine Size, Cylinders, Transmissions, dan CO Emissions. Kami berharap ini akan memberi kita gambaran lebih lengkap tentang faktor-faktor yang memengaruhi konsumsi bahan bakar kendaraan. Dengan menggunakan teknik data mining dan regresi linier berganda, penelitian ini bertujuan untuk mengembangkan model yang lebih akurat dalam memprediksi konsumsi bahan bakar, yang pada gilirannya dapat digunakan oleh produsen kendaraan, perusahaan transportasi, dan pembuat kebijakan untuk meningkatkan efisiensi bahan bakar, mengurangi biaya operasional, dan mendukung upaya pengurangan emisi gas rumah kaca. Analisis regresi sederhana memiliki maksud untuk mengetahui apakah suatu variabel dapat mempengaruhi variabel lainnya atau tidak (Sebastian Rudi, Pranoto, & Ariwibisono, 2023). Melalui penelitian ini, diharapkan program yang dibuat dapat menemukan hasil yang kemudian digunakan sebagai prediksi pembeli (Layli, Marhono, & Sari, 2025).

2. Metode

Penelitian ini merupakan jenis penelitian kuantitatif (Setiawan, Annur, & Melangi, 2024). Penelitian ini dikategorikan sebagai kuantitatif karena menghasilkan prediksi dan analisis terhadap data yang diperoleh dari Kaggle. Penelitian yang bertujuan untuk memprediksi dan menganalisis variabel-variabel dalam dataset dengan hasil berupa angka, termasuk dalam penelitian kuantitatif. Dalam hal ini, penelitian ini menggunakan metode Regresi Linear dan memanfaatkan 639 data dari 36 merek mobil. Studi penambangan data yang menggunakan berbagai lembar data dan model regresi linier telah dilakukan melalui berbagai aplikasi (Sholeh, Nurnawati, & Lestari, 2023). Regularized Linear Regression dipilih sebagai metode yang digunakan dalam penelitian ini untuk memprediksi jumlah konsumsi BBM berdasarkan variabel Vehicle Class, Engine Size, Cylinders, Transmissions, dan CO Emissions. Regresi linier adalah salah satu metode yang digunakan untuk meramalkan atau memprediksi suatu keluaran menggunakan dataset kualitatif maupun kuantitatif (Revaldi et al., 2024). Bahasa pemrograman yang digunakan dalam penelitian ini adalah Python, dan seluruh analisis data dilakukan melalui platform Kaggle. Simple Linear Regression juga merupakan pendekatan statistik yang digunakan untuk memprediksi karakteristik data baik kualitatif maupun kuantitatif (Alwy Yusuf et al., 2024). Analisis regresi merupakan salah satu teknik yang populer dan sering digunakan dalam penelitian data (Muttaqin & Srihartini, 2024). Regresi linier adalah metode statistika yang digunakan untuk membangun model hubungan antara variabel terikat (Y) dengan satu atau lebih variabel bebas (X) (Anggrawan, Hairani, & Azmi, 2022). Metode ini memanfaatkan garis lurus untuk menggambarkan hubungan antara dua variabel atau lebih. Variabel adalah besaran yang nilainya berubah-ubah, dan dikategorikan menjadi dua: variabel yang memengaruhi (penyebab) dan variabel yang terpengaruh (akibat). Prediksi dalam regresi linier dilakukan dengan asumsi bahwa pertumbuhan data bersifat linier, meskipun pada kenyataannya tidak selalu demikian (Jatikusumo & Hidayat, 2024). Regresi linier juga digunakan untuk menguji hubungan antara variabel yang memengaruhi dan variabel yang terpengaruh (Arief Hidayat, Ziyad, & Juliane, 2023).

3. Hasil Dan Pembahasan

Data yang terlampir merupakan hasil keluaran serta pengkajian menyeluruh dari kegiatan penelitian yang telah dilakukan. Penelitian ini menggunakan sebanyak 639 data kendaraan yang berasal dari berbagai merek mobil, antara lain Suzuki, Honda, Audi, Nissan, Ford, GMC, Mazda, Toyota, serta sejumlah brand lainnya yang turut diperhitungkan dalam analisis. Data tersebut mencakup berbagai varian kendaraan dengan spesifikasi dan karakteristik yang berbeda, sehingga memberikan cakupan yang luas dan representatif terhadap kondisi konsumsi bahan bakar pada kendaraan.

Dalam proses pengumpulan dan analisis data, penelitian ini memfokuskan diri pada beberapa variabel penting yang secara signifikan mempengaruhi efisiensi penggunaan bahan bakar kendaraan. Variabel-variabel tersebut meliputi model mesin dari masing-masing kendaraan, kelas mobil (seperti subcompact, compact, SUV, dll.), ukuran mesin (engine size) dalam satuan liter, jumlah silinder (cylinders), jenis

transmisi yang digunakan (manual atau otomatis), serta tingkat emisi karbon (CO Emissions) yang dihasilkan oleh kendaraan dalam satuan gram per kilometer.

Dengan memperhatikan kombinasi faktor-faktor tersebut, penelitian ini berupaya untuk membangun sebuah model prediktif yang mampu memperkirakan konsumsi bahan bakar secara lebih akurat menggunakan metode regresi linier regularisasi. Diharapkan bahwa hasil kajian ini dapat memberikan gambaran menyeluruh mengenai hubungan antara spesifikasi kendaraan dan efisiensi bahan bakarnya, serta menjadi dasar untuk pengambilan keputusan dalam pengembangan teknologi otomotif yang lebih ramah lingkungan dan hemat energi di masa depan.

TYPE	MODEL	VEHICLE CLASS	ENGINE SIZE	CYLINDERS	TRANSMISSION	COEMISSIONS
ACURA	1.6EL	COMPACT	1.6	4	A4	216
ACURA	1.6EL	COMPACT	1.6	4	M5	205
ACURA	3.2TL	MID-SIZE	3.2	6	A5	265
ACURA	3.5RL	MID-SIZE	3.5	6	A4	301
ACURA	INTEGRA	SUBCOMPACT	1.8	4	A4	230
ACURA	INTEGRA	SUBCOMPACT	1.8	4	M5	219
ACURA	INTEGRA GSR/TYPE R	SUBCOMPACT	1.8	4	M5	223
ACURA	NSX	SUBCOMPACT	3.0	6	A4	306
ACURA	NSX	SUBCOMPACT	3.2	6	M6	306
AUDI	A4	COMPACT	1.8	4	A5	251
AUDI	A4	COMPACT	1.8	4	M5	223
AUDI	A4	COMPACT	2.8	6	A5	285
AUDI	A4	COMPACT	2.8	6	M5	260
AUDI	A4 QUATTRO	COMPACT	1.8	4	A5	267
AUDI	A4 QUATTRO	COMPACT	1.8	4	M5	246
AUDI	A4 QUATTRO	COMPACT	2.8	6	A5	292
AUDI	A4 QUATTRO	COMPACT	2.8	6	M5	288
AUDI	A6	MID-SIZE	2.8	6	A5	294
AUDI	A6 AVANT QUATTRO	STATION WAGON - MID-SIZE	2.8	6	A5	306
AUDI	A6 QUATTRO	MID-SIZE	2.7	6	A5	304
AUDI	A6 QUATTRO	MID-SIZE	2.7	6	M6	304
AUDI	A6 QUATTRO	MID-SIZE	2.8	6	A5	306
AUDI	A6 QUATTRO	MID-SIZE	2.8	6	M5	297
AUDI	A6 QUATTRO	MID-SIZE	4.2	8	A5	308
AUDI	A8 QUATTRO	MID-SIZE	4.2	8	A5	308
AUDI	S4 QUATTRO	COMPACT	2.7	6	A5	304
AUDI	S4 QUATTRO	COMPACT	2.7	6	M6	304

Gambar 1.1. Hasil Kajian Penelitian

Berikut Data diatas adalah data yang belum diproses menggunakan metode Regularized Linear Regressions. Data yang ditampilkan diatas hanyalah sebagian, Karena datanya sangat banyak dan tidak memungkinkan untuk ditampilkan semuanya.

Setelah data diolah, data akan ditampilkan pada pembahasan ini. Namun penulis akan membahas terlebih dahulu data yang akan ditampilkan. Pada pembahasan sebelumnya penulis melibatkan beberapa brand mobil, model mobil, vehicle class, engine size, cylinder, transmisi, CO Emissions. Data yang telah diproses akan menampilkan Brand mobil, Vehicle class, Engine size, Cylinders, Transmisi, Co Emissions dan Predicted Fuel Consumption. Alasan penelitian ini melibatkan Brand adalah agar uji ini dapat dipercaya , karena brand yang digunakan juga hampir data semua brand mobil. Penulis juga melibatkan Vehicle Class, Karena berbedanya kelas kendaraan, tentu akan beda juga konsumsi bahan bakar minyaknya. Kemudian Mobil dengan cylinders yang berbeda juga akan membuat konsumsi bahan bakar yang berbeda juga. Kemudian Penulis juga menambahkan Transmisi sebagai pendukung ,

karena memang mobil sekarang transmisi ada 2 , yaitu manual dan automatic, penulis ingin mengetahui apakah ada perbedaan diantara 2 transmisi. Serta Penulis juga menambahkan hasil CO Emissions dari mobil mobil tersebut. Tujuannya adalah agar kedepannya mobil dapat diciptakan lebih rendah emisi karbon. Agar lingkungan dapat lebih terjaga dan dapat mengurangi polusi. Pada penelitian ini digunakan Data asli sebanyak 639 data , Jumlah data training 511, dan Jumlah data testing 128. Dapat didapatkan validasinya dari gambar dibawah ini.

```
# Cek jumlah data
print("Jumlah data asli:", len(X_new))
print("Jumlah data training:", len(X_train_new))
print("Jumlah data testing:", len(X_test_new))
```

Jumlah data asli: 639
 Jumlah data training: 511
 Jumlah data testing: 128

MAKE	VEHICLE CLASS	ENGINE SIZE	CYLINDERS	TRANSMISSION	COEMISSIONS	Predicted Fuel Consumption (L/100km)
ACURA	COMPACT	1.6	4	A4	216	9.46
ACURA	COMPACT	1.6	4	M5	205	9.15
ACURA	MID-SIZE	3.2	6	A5	265	12.19
ACURA	MID-SIZE	3.5	6	A4	301	12.18
ACURA	SUBCOMPACT	1.8	4	A4	230	9.17
ACURA	SUBCOMPACT	1.8	4	M5	219	8.86
ACURA	SUBCOMPACT	1.8	4	M5	223	8.86
ACURA	SUBCOMPACT	3.0	6	AS4	306	11.96
ACURA	SUBCOMPACT	3.2	6	M6	306	11.93
AUDI	COMPACT	1.8	4	A5	251	9.53
AUDI	COMPACT	1.8	4	M5	223	9.3
AUDI	COMPACT	2.8	6	A5	285	11.41
AUDI	COMPACT	2.8	6	M5	260	11.18
AUDI	COMPACT	1.8	4	A5	267	9.53
AUDI	COMPACT	1.8	4	M5	246	9.3
AUDI	COMPACT	2.8	6	A5	292	11.41
AUDI	COMPACT	2.8	6	M5	288	11.18
AUDI	MID-SIZE	2.8	6	A5	294	11.56
AUDI	STATION WAGON - MID-SIZE	2.8	6	A5	306	11.93
AUDI	MID-SIZE	2.7	6	A5	304	11.49
AUDI	MID-SIZE	2.7	6	M6	304	12.14
AUDI	MID-SIZE	2.8	6	A5	306	11.56
AUDI	MID-SIZE	2.8	6	M5	297	11.34
AUDI	MID-SIZE	4.2	8	A5	308	13.75
AUDI	MID-SIZE	4.2	8	A5	308	13.75
AUDI	COMPACT	2.7	6	A5	304	11.33
AUDI	COMPACT	2.7	6	M6	304	11.98
AUDI	MINICOMPACT	1.8	4	M5	260	9.8
AUDI	MINICOMPACT	1.8	4	M6	260	10.67

Gambar 1.2. Hasil Proses Penelitian

Data yang diprediksi juga menampilkan jumlah konsumsi bahan bakar per 100km . Data yang ditampilkan juga hanya Sebagian dan tidak semua. Karena tidak memungkinkan untuk di tampilkan semua. Penelitian ini menggunakan Metode Regualized Linear Regressions. Aplikasi coding yang dipakai untuk penelitian ini adalah KAGGLE dan dengan Bahasa pemograman Phyton.

```
[11]: # Menampilkan hasil evaluasi model
evaluation_new_results = {
    "Mean Absolute Error (MAE)": round(mae_new, 2),
    "Mean Squared Error (MSE)": round(mse_new, 2),
    "Root Mean Squared Error (RMSE)": round(rmse_new, 2),
    "R-squared (R2) Score": round(r2_new, 2)
}

print(evaluation_new_results)

{'Mean Absolute Error (MAE)': 0.92, 'Mean Squared Error (MSE)': 2.01, 'Root Mean Squared Error (RMSE)': 1.42, 'R-squared (R2) Score': 0.79}
```

Gambar 1.3. Hasil Akhir Penelitian

Dapat Dilihat pada gambar diatas, penelitian ini menghasilkan MAE 0.92, MSE 2.01, RMSE 1.42 , R-SQUARES 0.79 . Dapat disimpulkan Nilai Mean Absolute Error (MAE) = 0.92 dan Root Mean Squared Error (RMSE) = 1.42 menunjukkan bahwa model memiliki kesalahan prediksi yang relatif rendah. Artinya, prediksi model cukup dekat dengan nilai sebenarnya. Nilai R-squared (R^2) = 0.79 menandakan bahwa model dapat menguraikan 79% variasi dalam data. Hasil tersebut menunjukkan bahwa model yang dipakai cukup baik untuk mencakup pola dan relasi yang ada pada data. Dan dapat disimpulkan untuk metode ini 79% nilai Akurasinya. Nilai Mean Squared Error (MSE) = 2.01 mengindikasikan bahwa kesalahan kuadrat antara prediksi dan nilai sebenarnya masih berada pada level yang wajar. MSE lebih sensitif terhadap kesalahan yang lebih besar, namun nilainya tidak menunjukkan adanya kesalahan besar yang signifikan.

4. Simpulan

Berdasarkan kegiatan yang dilakukan, dapat disimpulkan bahwa penerapan metode Regularized Linear Regressions untuk memprediksi penggunaan bahan bakar minyak memberikan hasil yang cukup baik dan akurat. Nilai koefisien determinasi (R) yang diperoleh sebesar 0,79 menunjukkan bahwa model ini mampu menjelaskan sekitar 79% variasi data, yang berarti akurasinya cukup tinggi karena nilai R mendekati 1. Nilai R yang rendah, seperti di bawah 0,5, akan menunjukkan bahwa metode ini kurang cocok, namun dalam penelitian ini hasilnya menunjukkan kinerja yang memuaskan. Selain itu, dari 639 data mobil dari berbagai merek yang diteliti, dapat disimpulkan bahwa mobil yang paling hemat bahan bakar adalah dari merek Chevrolet dengan kelas mobil Sub Compact, ukuran mesin 1.0, 3 silinder, transmisi manual, emisi karbon sebesar 147, dan konsumsi bahan bakar 7.69 liter per 100 km. Di posisi berikutnya ada merek Honda dengan spesifikasi yang sama namun emisi karbon lebih rendah yaitu 104, dan konsumsi bahan bakar yang sama. Terakhir, merek Pontiac juga memiliki efisiensi serupa dengan kelas Sub Compact, mesin 1.0, 3 silinder, transmisi manual, yang turut menunjukkan konsumsi bahan bakar rendah.

5. Daftar Pustaka

- Anggrawan, A., Hairani, H., & Azmi, N. (2022). Prediksi penjualan produk Unilever menggunakan metode regresi linear. *Jurnal Bumigora Information Technology (BITe)*, 4(2), 123–132. <https://doi.org/10.30812/bite.v4i2.2416>
- Augie Sugiarto Nunka, & Pranoto, W. J. (2024). Metode regresi linier berganda untuk prediksi pemakaian BBM PT. Kalonica Bara Kusuma. *Jupiter: Publikasi Ilmu*

- Keteknikan Industri, Teknik Elektro dan Informatika*, 2(1), 78–90.
<https://doi.org/10.61132/jupiter.v2i1.56>
- Deni, M., & Latifah, D. R. (2021). Prediksi pengisian BBM HSD dengan metode multiple linear regression. *Jurnal Sistem Informasi dan Teknologi (JUST-IT)*, 11(3).
<https://jurnal.umj.ac.id/index.php/just-it/index>
- Hidayat, A., Ziyad, M. T., & Juliane, C. (2023). Memprediksi volume sampah di Jawa Barat dengan metode regresi linier. *Jurnal Teknik Informatika dan Sistem Informasi*.
- Husdi, H., & Dalai, H. (2023). Penerapan metode regresi linear untuk prediksi jumlah bahan baku produksi selai Bilfagi. *Jurnal Informatika*, 10(2), 129–135.
<https://doi.org/10.31294/inf.v10i2.14129>
- Indarwati, T., Irawati, T., & Rimawati, E. (2019). Penggunaan metode linear regression untuk prediksi penjualan smartphone. *Jurnal Teknologi Informasi dan Komunikasi (TIKomsin)*, 6(2). <https://doi.org/10.30646/tikomsin.v6i2.369>
- Jatikusumo, D., & Hidayat, R. R. (2024). Optimasi penentuan lokasi bencana alam dengan regresi linier sederhana dan berganda. *Jurnal Informatika dan Teknik Elektro Terapan*, 12(3S1). <https://doi.org/10.23960/jitet.v12i3S1.5257>
- Kurniawan, I., Rahaningsih, N., & Suprpti, T. (2024). Implementasi algoritma regresi linier dan K-nearest neighbor untuk prediksi harga rumah. *Jurnal Mahasiswa Teknik Informatika*, 8(1).
- Layli, I. N., Marhono, A. P., & Sari, A. P. (2025). Prediksi jumlah pembeli pulsa berdasarkan jumlah pembeli dan total harga dengan metode regresi linier. *Prosiding Seminar Implementasi Teknologi Informasi dan Komunikasi*, 4(1).
<https://doi.org/10.31284/p.semtik.2025-1.6946>
- Linear, P. R., Prediksi, U., Beras, H., Indonesia, D., Arinal, V., & Azhari, M. (2023). Penerapan regresi linear untuk prediksi harga beras di Indonesia. *Jurnal Sains dan Teknologi*, 5(1). <https://doi.org/10.55338/saintek.v5i1.1417>
- Muttaqin, Z., & Srihartini, E. (2024). Penerapan metode regresi linier sederhana untuk prediksi persediaan obat jenis tablet. *Jurnal Sistem Informasi*, 9(1), 12–16.
- Revaldi, K. A., Cahyono, F. D., Hakimah, M., Rotul, R., Institut, M., Adhi, T., & Surabaya, T. (2024). Penerapan metode regresi linier sederhana dalam memprediksi jumlah kebutuhan ekspor migas dan non-migas di Indonesia. *Prosiding SNESTIK - Seminar Nasional Teknik Elektro, Sistem Informasi, dan Teknik Informatika*. <https://doi.org/10.31284/p.snestik.2023.4246>
- Rudi, S., Pranoto, Y. A., & Ariwibisono, F. X. (2023). Penerapan metode regresi linier dalam peramalan penjualan kue di Toko Karya Bahari Samarinda berbasis website. *Jurnal Mahasiswa Teknik Informatika*, 7(4).
- Setiawan, A., Annur, H., & Melangi, S. (2024). Penerapan metode regresi linier sederhana untuk prediksi jumlah persediaan pestisida. *Jurnal Ilmiah Ilmu Komputer Banthayo Lo Komputer*, 3(1).
- Sholeh, M., Nurnawati, E. K., & Lestari, U. (2023). Penerapan data mining dengan metode regresi linear untuk memprediksi data nilai hasil ujian menggunakan RapidMiner. *Jurnal Informatika Sunan Kalijaga*, 8(1).
<https://archive.ics.uci.edu/ml/datasheets.php>
- Suwaryo, N., Rahman, A., Marini, D., Atmaja, U., & Basri, A. (2023). Prediksi penyakit diabetes untuk pencegahan dini dengan metode regresi linear. *Bulletin of*

Information Technology (BIT), 4(2), 313–319.
<https://doi.org/10.47065/bit.v3i1>

Yusuf, M. A., Abraham, A., & Rukmana, H. (2024). Analisis regresi linier sederhana dan berganda beserta penerapannya. *Journal on Education*, 6(2).