

Peramalan Jumlah Kendaraan Diuji Baik Menggunakan Metode Regresi Linier Berganda

Agus Astra Bawa^{1*}, I Nyoman Sukajaya¹, I Made Agus Wirawan¹

¹Program Studi Ilmu Komputer, Universitas Pendidikan Ganesha, Indonesia.

Artikel Info

Kata Kunci:

Mean Absolut Percentage Error;
Peramalan;
Regresi Linier Berganda.

Keywords:

Mean Absolute Percentage Error;
Forecasting;
Multiple Linear Regression.

Riwayat Artikel:

Submitted: 16 Maret 2025

Accepted: 30 Maret 2025

Published: 31 Maret 2025

Abstrak: Adanya permasalahan dalam penyediaan bahan baku pengujian di Dinas Perhubungan Kabupaten Buleleng sangat erat kaitannya dengan belum adanya peramalan jumlah kendaraan diuji baik. Peramalan juga tidak bisa dilakukan tanpa teridentifikasinya variabel-variabel yang mempengaruhi jumlah kendaraan diuji baik tersebut. Tujuan penelitian ini adalah untuk menentukan variabel yang tepat yang mempengaruhi jumlah kendaraan diuji baik serta menghitung hasil peramalan jumlah kendaraan diuji baik di Dinas Perhubungan Kabupaten Buleleng. Metode analisis prediksi menggunakan Regresi Linier Berganda. Evaluasi peramalan menggunakan Mean Absolute Percentage Error (MAPE). Penelitian ini menghasilkan 5 variabel bebas yaitu jumlah hari kerja dibulan bersangkutan, usia kendaraan dari kendaraan diuji baik 6 bulan sebelumnya, jarak tempuh (alamat pemilik) dari kendaraan diuji baik 6 bulan sebelumnya, jenis kendaraan mobil barang dari kendaraan diuji baik 6 bulan sebelumnya, dan sifat kendaraan umum dari kendaraan diuji baik 6 bulan sebelumnya. Dengan persamaan Regresi Linier Berganda $Y = 773,884 + 11,695X_1 - 23,788X_3 - 11,633X_4 + 0,005X_5 + 1,502X_8$ menghasilkan peramalan jumlah kendaraan diuji baik dari bulan Januari 2024 sampai dengan bulan Juni 2024 adalah 600, 540, 595, 577, 570 dan 581 dengan hasil uji akurasi MAPE sebesar 5,39 dengan kategori sangat bagus.

Abstract: The existence of problems in the provision of raw materials for testing at the Buleleng Regency Transportation Agency is closely related to the absence of a forecast of the number of vehicles tested properly. Forecasting also cannot be done without identifying the variables that influence the number of vehicles tested properly. The purpose of this study was to determine the right variables that influence the number of vehicles tested properly and to calculate the results of the forecast of the number of vehicles tested properly at the Buleleng Regency Transportation Agency. The prediction analysis method uses Multiple Linear Regression. Forecast evaluation uses Mean Absolute Percentage Error (MAPE). This study produced 5 independent variables, namely the number of working days in the relevant month, the age of the vehicle from the vehicle tested properly 6 months earlier, the distance traveled (owner's address) from the vehicle tested properly 6 months earlier, the type of goods vehicle from the vehicle tested properly 6 months earlier, and the nature of public vehicles from the vehicle tested properly 6 months earlier. With the Multiple Linear Regression equation $Y = 773.884 + 11.695X_1 - 23.788X_3 - 11.633X_4 + 0.005X_5 + 1.502X_8$, the forecast for the number of vehicles tested from January 2024 to June 2024 is 600, 540, 595, 577, 570 and 581 with a MAPE accuracy test result of 5.39 with a very good category.

Corresponding Author:

Agus Astra Bawa

Email: agus.astra@undiksha.ac.id

PENDAHULUAN

Pelayanan publik adalah aktivitas yang bertujuan untuk memenuhi kebutuhan layanan sesuai dengan ketentuan hukum yang berlaku bagi setiap warga negara dan penduduk. Layanan ini mencakup penyediaan barang, jasa, serta pelayanan administrasi yang diselenggarakan oleh pihak penyedia layanan publik (Utami et al., 2020). Pengujian kendaraan bermotor (PKB) meliputi serangkaian aktivitas evaluasi dan inspeksi terhadap komponen-komponen kendaraan bermotor, termasuk kereta gandengan, kereta tempelan, dan kendaraan khusus (Rahma & Sihombing, 2023). Tujuan utamanya adalah memastikan kepatuhan terhadap standar teknis dan kelayakan jalan yang telah ditetapkan.

Sampai saat ini, pengadaan bahan baku PKB pada Dinas Perhubungan Kabupaten Buleleng (Dishub Buleleng) masih belum menggunakan dasar perencanaan yang pasti. Permasalahan dalam pengadaan bahan baku PKB sangat erat kaitannya dengan jumlah kendaraan diuji baik yang belum terprediksi. Kedatangan jumlah kendaraan diuji baik juga sangat berkaitan erat dengan adanya variabel-variabel lain yang mempengaruhi. Oleh karena itu, penulis berupaya mengidentifikasi variabel-variabel yang mempengaruhi jumlah kendaraan diuji baik tersebut dan kemudian melakukan peramalan terhadap jumlah kendaraan diuji baik periode tertentu di PKB Dishub Buleleng.

Peramalan adalah memperkirakan besar atau jumlah dari sesuatu pada masa mendatang menggunakan data yang sudah ada pada masa sebelumnya kemudian dianalisis dengan cara alamiah dan cara khusus menggunakan metode *statistic* (Muhammad Nur et al., 2024). Pentingnya peramalan terletak pada fakta bahwa peramalan menjadi titik awal dari segala perencanaan. Artinya, kualitas perencanaan sangat bergantung pada kualitas peramalan (Mira tania et al., 2024). Terdapat dua jenis metode peramalan, yaitu metode peramalan kuantitatif dan metode peramalan kualitatif. Metode kuantitatif adalah metode peramalan yang diterapkan ketika data historis tersedia untuk digunakan sebagai dasar analisis. Metode kuantitatif terbagi menjadi tiga kategori, yaitu metode deret waktu, metode kausal, serta kombinasi dari metode deret waktu dan metode kausal (Susilawati & Sunendiari, 2022). Metode kausal mencoba mengidentifikasi faktor-faktor yang memengaruhi variabel yang ingin diramalkan dan memproyeksikan variabel tersebut berdasarkan faktor-faktor lain yang mempengaruhinya. Metode kausal yang paling umum adalah metode regresi. Beberapa metode peramalan yang lain meliputi metode *K-Means*, metode Logika *Fuzzy*, metode ARIMA, dan metode lainnya.

Dalam beberapa penelitian yang relevan sebelumnya, metode Regresi Linier Berganda lebih unggul dibandingkan metode lain karena memiliki tingkat kesalahan yang lebih rendah, seperti pada penelitian Annisa Rismanitanti (2022) menggunakan *K-Means* dalam studi kasus Perbandingan Pengolahan Data Prediksi Persediaan Gas LPG 3 Kg Pada PT. Blora Mustika Kotabumi Menggunakan Regresi Linier Berganda dan *K-Means* dimana hasil peramalan menunjukkan nilai *error* metode Regresi Linier Berganda lebih kecil dibandingkan nilai *error* metode *K-Means* sehingga disimpulkan metode RLB lebih lebih baik dibandingkan dengan metode *K-Means*. Demikian juga pada penelitian (Yulia Retni Sari (2021) menggunakan metode Regresi Linier Berganda dan Logika *Fuzzy* dalam studi kasus Perbandingan Analisis Logika *Fuzzy* dan Regresi Linier Berganda Dalam Menentukan Produksi Beras Nasional dimana hasil peramalan berdasarkan nilai kesalahan relatif sebagai nilai uji akurasi *error* menghasilkan metode Regresi Linier Berganda lebih baik dari metode Logika *Fuzzy*. Penelitian relevan yang lain dilakukan oleh Hilman Winnos (2022) dengan judul Perbandingan Metode Regresi Linier Berganda dan *Autoregressive Integrated Moving Average* (ARIMA) Untuk Prediksi Saham PT. BSI, Tbk, penelitian ini memiliki tujuan untuk menerapkan dan mengembangkan model persamaan yang dapat memberikan prediksi harga saham PT. BSI Tbk. dengan akurasi tinggi, membandingkan dua metode yang digunakan. Hasil penelitian menunjukkan bahwa model Regresi Linier Berganda berhasil

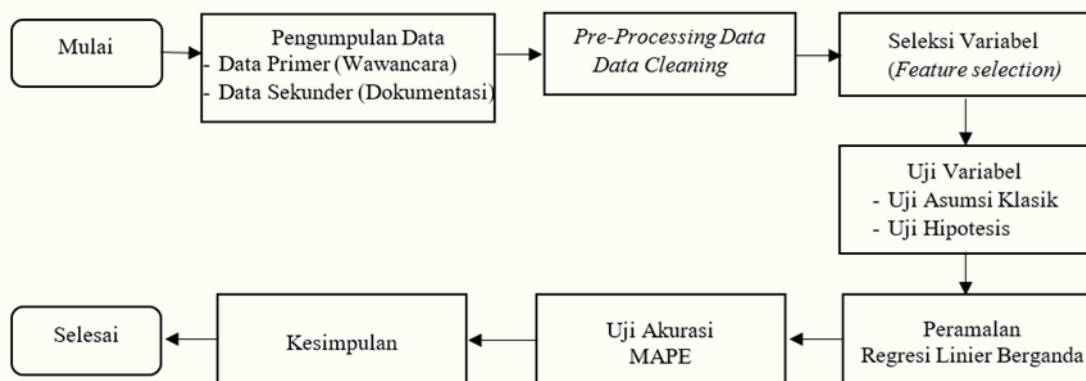
menghasilkan nilai MAPE sebesar 1,1%, menunjukkan tingkat akurasi sebesar 98,9%. Hal ini menunjukkan bahwa model tersebut lebih akurat dibandingkan dengan model ARIMA, yang menghasilkan nilai MAPE sebesar 2,36% dan tingkat akurasi sebesar 98,9%. Mengacu latar belakang dan permasalahan di atas, penulis menggunakan metode kausal Regresi Linier dalam melakukan penelitian. Adanya variabel bebas yang kemungkinan lebih dari 1 (satu) sehingga penulis menggunakan Metode Regresi Linier Berganda.

Dalam metode kausal, variabel terikat sangat ditentukan oleh variabel bebasnya. Keakuratan nilai variabel terikat sangat bergantung dari ketepatan penentuan variabel bebasnya. Tujuan penelitian ini adalah untuk menentukan variabel yang tepat yang mempengaruhi jumlah kendaraan diuji baik serta menghitung hasil peramalan jumlah kendaraan diuji baik di Dinas Perhubungan Kabupaten Buleleng. Penentuan variabel yang tepat dilakukan dengan seleksi variabel dan uji variabel untuk selanjutnya digunakan dalam meramalkan jumlah kendaraan diuji baik di Dishub Buleleng menggunakan metode Regresi Linier Berganda dengan uji akurasi *Mean Absolute Percentage Error* (MAPE).

METODE

Penelitian ini adalah kombinasi antara penelitian kualitatif dan kuantitatif. Penelitian kualitatif dilakukan dengan metode wawancara dengan ahli PKB dan penelitian kuantitatif dilakukan dengan memanfaatkan data sekunder yang diperoleh dari Dishub Buleleng.

Rancangan penelitian merujuk pada keseluruhan perencanaan suatu penelitian, yang melibatkan aspek-aspek seperti perumusan hipotesis, analisis data, dan langkah-langkah hingga tahap akhir. Rancangan penelitian ini dibuat dengan tujuan untuk memperlihatkan rencana pemecahan masalah peramalan jumlah kendaraan diuji baik pada Dishub Buleleng dengan metode Regresi Linier Berganda, dapat dilihat pada Gambar 1.



Gambar 1. Rancangan Penelitian

Proses pengumpulan data terdiri dari dua jenis, yaitu data primer dan data sekunder. Data primer dikumpulkan melalui metode wawancara dengan seorang ahli PKB. Wawancara ahli ini dilakukan sebagai dasar dalam penentuan variable apa saja yang mempengaruhi jumlah kendaraan diuji baik, mengingat dalam penelitian-penelitian sebelumnya, belum ditemukan tentang adanya topik yang sama terkait variabel yang mempengaruhi jumlah kendaraan diuji baik. Pengumpulan data sekunder dilakukan dengan meminta data jumlah kendaraan yang diuji baik pada periode bulan Januari 2022 hingga bulan Desember 2023 di unit PKB Dishub Buleleng.

Sebelum melakukan proses penghitungan, dilakukan *pre-processing data* untuk menyiapkan data sesuai dengan kebutuhan pada proses penghitungannya. *Pre-Processing Data* dilakukan dengan proses pembersihan data. *Data cleaning* (pembersihan data) adalah prosedur untuk memastikan bahwa data dalam dataset terjamin kebenarannya, konsistensinya, dan kegunaannya (Algoritma, 2022). Pembersihan Data (*data cleaning*) diperlukan untuk menghapus data yang tidak representatif, seperti

outlier, serta data yang tidak lengkap atau kosong. Data dengan kualitas rendah, seperti data yang tidak akurat, tidak lengkap, atau tidak konsisten, dapat menyebabkan kesimpulan yang bias, model yang kurang andal, atau interpretasi yang keliru (Hanafi, 2023).

Seleksi Variabel (*Feature selection*) adalah proses mengurangi jumlah fitur atau variabel input dengan memilih fitur-fitur yang dianggap paling relevan terhadap model (Urfah, 2022). Seleksi variabel menggunakan metode *Filter Feature Selection (correlation coefficient)* pada tools WEKA (*Waikato Environment for Knowledge Analysis*). *Correlation Attribute Evaluation* merupakan metode pemilihan fitur yang menggunakan metode pencarian peringkat (Allo, 2023). Metode ini memilih fitur-fitur yang relevan tanpa tergantung pada jenis algoritma pembelajaran yang digunakan. Pendekatan ini menggunakan pengukuran statistik untuk menilai relevansi setiap fitur. Metode ini efisien secara komputasi dan berjalan dengan cepat. Konsep kerjanya menggunakan metode statistika untuk menilai atributnya, kemudian setiap atribut akan diurutkan sesuai tingkatannya serta dibuat matriks guna mengetahui atribut yang relevan, sedangkan atribut yang tidak relevan akan dieliminasi (Azizah et al., 2023). *Threshold* (ambang batas) merupakan nilai batas relevan untuk pemilihan fitur optimal. Nilai threshold berada pada interval 0 sampai 1 dan penggunaannya bersifat independen (tergantung pada pengguna).

Uji Variabel ditentukan dengan Uji Asumsi Klasik dan Uji Hipotesis. Untuk memperoleh persamaan regresi yang optimal, dilakukan pengujian asumsi klasik yang meliputi uji normalitas residual, uji multikolinieritas, uji heteroskedastisitas, uji autokorelasi, uji t, uji f, dan analisis determinasi. Uji Asumsi Klasik merupakan proses analisis yang digunakan untuk mengevaluasi apakah suatu model regresi linear *Ordinary Least Square* (OLS) memenuhi asumsi-asumsi dasar yang diperlukan atau terdapat pelanggaran terhadap asumsi-asumsi tersebut (Mardiatmoko, 2020). Analisa Variabel diawali dengan syarat dalam melakukan perhitungan Regresi Linier Berganda yaitu melalui Uji Asumsi Klasik. Uji asumsi klasik dilakukan sebelum pengujian hipotesis untuk memverifikasi apakah model regresi memenuhi persyaratan ekonometrik. Ini melibatkan pengujian normalitas, multikolinearitas, heteroskedastisitas, dan autokorelasi guna memastikan bahwa persamaan regresi dapat diterima secara statistik (Purba et al., 2021). Oleh karena itu, setiap model yang telah dikembangkan perlu mengalami uji residu untuk memastikan bahwa asumsi-asumsi ini telah terpenuhi. Jika asumsi-asumsi tersebut terpenuhi, maka model regresi tersebut dapat dianggap dapat digunakan. Sebaliknya, jika terdapat pelanggaran terhadap asumsi atau ketidakpenuhannya, perlu dilakukan penanganan lebih serius karena penyimpangan dari asumsi-asumsi tersebut dapat mengakibatkan kesalahan dalam melakukan penarikan kesimpulan, atau bahkan dapat membuat model regresi kurang efektif dalam melakukan estimasi terhadap rata-rata respons. Pelaksanaan uji asumsi klasik dan uji hipotesis menggunakan aplikasi SPSS. SPSS (*Statistical Package for the Social Sciences*) adalah perangkat lunak statistik yang dirancang untuk melakukan analisis data tingkat lanjut, termasuk pengolahan data menggunakan algoritma *machine learning*, *analisis string*, dan *analisis big data* yang terintegrasi untuk mendukung pengembangan platform analisis data (Ramadhani et al., 2023). Aplikasi SPSS memainkan peran penting dalam menyederhanakan analisis data asumsi klasik dalam penelitian. Sebagai perangkat lunak statistik yang banyak digunakan, SPSS menawarkan kemudahan dan efisiensi dalam menguji asumsi-asumsi statistik yang menjadi dasar analisis data (Sabtohadhi et al., 2024).

Uji Asumsi Klasik terdiri dari (Sari, 2022) :

1. Uji Normalitas

Tujuan : untuk menilai apakah distribusi residual bersifat normal

2. Uji Multikolinieritas

Tujuan : Untuk menentukan apakah terdapat hubungan antara variabel bebas dalam model regresi

3. Uji Heteroskedastisitas

Tujuan : Untuk menentukan apakah terjadi variasi yang tidak konsisten dari residu

4. Uji Autokorelasi

Tujuan : untuk menguji ada atau tidaknya korelasi pada residual pada periode t dengan residual pada periode sebelumnya (t-1)

Asumsi Hipotesis terdiri dari :

1) Uji-T (Pengaruh Variabel Independent secara parsial terhadap Variabel Dependent)

Uji t, yang juga dikenal sebagai uji parsial, digunakan untuk mengevaluasi dampak individual masing-masing variabel independen terhadap variabel dependennya.

2) Uji-F (Pengaruh Variabel Independent secara simultan terhadap Variabel Dependent)

Uji F, juga dikenal sebagai Uji Serentak atau Uji Model/Uji Anova, merupakan metode untuk mengevaluasi pengaruh bersama-sama dari semua variabel bebas terhadap variabel terikat dalam suatu model.

3) Koefisien Determinasi

Pengujian koefisien determinasi bertujuan untuk mengevaluasi sejauh mana model mampu menjelaskan pengaruh bersama dari variabel independen terhadap variabel dependen, yang diwakili oleh nilai *adjusted R-Squared*.

Kausalitas mengacu pada hubungan sebab-akibat antara dua atau lebih variabel. Pengukuran ini bertujuan untuk menentukan sejauh mana faktor penyebab memengaruhi terjadinya atau terbentuknya variabel akibat (Ruspendi et al., 2022). Tujuan dari metode ini adalah untuk menemukan suatu fungsi yang dapat menjelaskan ketergantungan variabel tak bebas terhadap semua variabel bebas. Metode ini dikenal sebagai regresi linier berganda.

Menurut Y. Asohi and A. Andri, Regresi Linier Berganda adalah metode untuk memprediksi dampak dari dua atau lebih variabel prediktor terhadap satu *variabel respons*, dengan tujuan untuk menguji keberadaan hubungan fungsional antara variabel independen (X) dan variabel dependen (Y) (Rismanitanti et al., 2022). Analisis regresi linier berganda bertujuan untuk melakukan prediksi nilai *variabel respons* (Y) berdasarkan nilai-nilai variabel prediktor (X1, X2, ..., Xn) yang diketahui. Selain itu, analisis ini juga membantu dalam memahami arah dan kekuatan hubungan antara variabel respons dan variabel prediktor tersebut. Metode ini juga berguna untuk memperoleh hasil dari suatu hipotesis (Marta et al., 2023). Hasil dari metode ini berupa persamaan garis dengan koefisien yang mencerminkan kecenderungan nilai, baik itu positif maupun negatif. Persamaan Regresi Linier Berganda ditunjukkan pada Persamaan (1) berikut (Trunfio et al., 2022).

$$Y = a + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n \dots \dots \dots (1)$$

Dari hasil peramalan, kemudian dihitung tingkat akurasi. Untuk mengevaluasi kesalahan dalam model peramalan, dapat digunakan ukuran seperti *Mean Absolute Percentage Error* (MAPE) (Rohaeti et al., 2022). MAPE dihitung dengan menambahkan semua error absolut pada setiap posisi data point (perbedaan antara nilai aktual dan nilai prediksi), kemudian membaginya dengan nilai aktual pada posisi tersebut. Selanjutnya, dihitung rata-rata dari kesalahan persentase absolut tersebut. Hasilnya kemudian dikonversi menjadi persentase dengan mengalikan dengan 100%. Berikut adalah perhitungan berdasarkan rumus MAPE, ditunjukkan pada Persamaan (2) berikut (Ahmad et al., 2022).

$$MAPE = \left(\frac{1}{N} \sum_{t=1}^n \frac{|y_t - \hat{y}_t|}{y_t} \right) \times 100\% \dots \dots \dots (2)$$

HASIL DAN PEMBAHASAN

Pengumpulan data primer dilakukan dengan metode wawancara dengan seorang ahli PKB. Wawancara ahli ini dilakukan sebagai dasar dalam penentuan variabel apa saja yang mempengaruhi jumlah kendaraan diuji baik (Y), mengingat dalam penelitian-penelitian sebelumnya, belum ditemukan tentang adanya topik yang sama terkait variabel yang mempengaruhi jumlah kendaraan diuji baik.

Pengumpulan data sekunder dilakukan dengan meminta data jumlah kendaraan yang diuji baik pada periode bulan Januari 2022 hingga bulan Desember 2023 di unit PKB Dishub Buleleng. Pengumpulan data primer dan data sekunder menghasilkan data variabel bebas sebagai berikut :

- 1 Jumlah Hari Kerja dibulan bersangkutan sebagai variabel bebas 1 (X_1),
- 2 Jumlah Kendaraan Diuji Baik 6 bulan sebelumnya sebagai variabel bebas 2 (X_2),
- 3 Usia Kendaraan 6 bulan sebelumnya sebagai variabel bebas 3 (X_3),
- 4 Jarak Tempuh/Alamat 6 bulan sebelumnya sebagai variabel bebas 4 (X_4),
- 5 Jenis Kendaraan Mobil Barang 6 bulan sebelumnya sebagai variabel bebas 5 (X_5),
- 6 Jenis Kendaraan Mobil Penumpang 6 bulan sebelumnya sebagai variabel bebas 6 (X_6),
- 7 Sifat Kendaraan Tidak Umum 6 bulan sebelumnya sebagai variabel bebas 7 (X_7) dan
- 8 Sifat Kendaraan Umum 6 bulan sebelumnya sebagai variabel bebas 8 (X_8).

Pada data sekunder jumlah kendaraan diuji baik dari bulan Januari 2022 sampai dengan bulan Desember 2023, terdapat 16.078 data yang dilakukan proses data cleaning. Hasilnya terdapat beberapa data yang kosong/missing dan juga data outlier karena kesalahan pengetikan pada saat penginputan pada variabel tahun kendaraan maupun alamat (jarak tempuh). Data-data yang kosong maupun outlier tersebut kemudian dilengkapi dan diperbaiki, dirata-ratakan menjadi data bulanan untuk menghasilkan data uji yang siap untuk dilakukan perhitungan seperti ditunjukkan pada Tabel 1.

Tabel 1. Hasil *Data Cleaning*

No	Bulan	Y	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8
1	Juli 2022	787	21	784	11	11,76	767	17	683	101
2	Agustus 2022	783	22	648	12	12,58	624	24	552	96
3	September 2022	770	22	754	12	12,23	736	18	638	116
4	Oktober 2022	610	20	749	15	17,49	722	27	650	99
5	November 2022	685	22	662	14	14,97	638	24	575	87
6	Desember 2022	735	22	834	12	12,85	811	23	734	100
7	Januari 2023	648	16	787	14	12,71	765	22	680	107
8	Februari 2023	659	20	783	14	14,33	756	27	704	79
9	Maret 2023	670	20	770	13	14,48	743	27	663	107
10	April 2023	501	14	610	15	19,62	592	18	519	91
11	Mei 2023	599	20	685	14	15,29	666	19	601	84
12	Juni 2023	603	17	735	14	16,81	717	18	647	88
13	Juli 2023	629	20	648	13	15,34	618	30	562	86
14	Agustus 2023	654	18	659	13	13,47	641	18	572	87
15	September 2023	619	20	670	14	15,81	643	27	582	88
16	Oktober 2023	601	22	501	16	15,49	481	20	436	65
17	November 2023	566	22	599	15	15,29	578	21	529	70
18	Desember 2023	528	18	603	15	15,40	579	24	528	75

Seleksi variabel menggunakan metode *Filter Feature Selection (correlation coefficient)* pada tools WEKA. Metode ini melakukan pemilihan fitur yang relevan tanpa bergantung pada jenis algoritma pembelajaran yang diterapkan. Pendekatan ini menggunakan pengukuran statistik untuk menilai relevansi setiap fitur. Metode ini efisien secara komputasi dan berjalan dengan cepat. Hasil seleksi variabel ditunjukkan pada Gambar 2 berikut.

```

=== Attribute Selection on all input data ===

Search Method:
    Attribute ranking.

Attribute Evaluator (supervised, Class (numeric): 1 Jumlah Kendaraan Diuji Baik
    Correlation Ranking Filter

Ranked attributes:
0.5992  9 Sifat Kendaraan Umum                (X8)
0.5783  2 Jumlah Hari Kerja (Hari/Bulan)        (X1)
0.5458  6 Jenis Kendaraan Mobil Barang          (X5)
0.5421  3 Jumlah Kendaraan Diuji Baik           (X2)
0.5037  8 Sifat Kendaraan Tidak Umum            (X7)
-0.0427 7 Jenis Kendaraan Mobil Penumpang      (X6)
-0.8354 5 Jarak Tempuh/Alamat                  (X4)
-0.8655 4 Usia Kendaraan                       (X3)
    
```

Gambar 2. Hasil Seleksi Variabel

Threshold (ambang batas) adalah nilai yang digunakan untuk menentukan relevansi dalam pemilihan fitur optimal. Nilai ini berada dalam rentang antara 0 hingga 1 dan penggunaannya bersifat independen (tergantung pada pengguna) (Kesuma, 2011). Dari 8 variabel bebas yang dilakukan seleksi, terlihat pengaruh dari variabel X6 memiliki nilai korelasi sangat kecil sehingga untuk proses selanjutnya, variabel bebas ini akan dihilangkan. Hasil seleksi variabel menghasilkan 7 variabel bebas yaitu X₁, X₂, X₃, X₄, X₅, X₇ dan X₈.

Uji Variabel ditentukan dengan Uji Asumsi Klasik dan Uji Hipotesis. Uji Variabel dilakukan pada data seperti ditunjukkan pada Tabel 2 berikut.

Tabel 2. Data Uji Variabel

No	Bulan	Y	X ₁	X ₂	X ₃	X ₄	X ₅	X ₇	X ₈
1	Juli 2022	787	21	784	11	11,76	767	683	101
2	Agustus 2022	783	22	648	12	12,58	624	552	96
3	September 2022	770	22	754	12	12,23	736	638	116
4	Oktober 2022	610	20	749	15	17,49	722	650	99
5	November 2022	685	22	662	14	14,97	638	575	87
6	Desember 2022	735	22	834	12	12,85	811	734	100
7	Januari 2023	648	16	787	14	12,71	765	680	107
8	Februari 2023	659	20	783	14	14,33	756	704	79
9	Maret 2023	670	20	770	13	14,48	743	663	107
10	April 2023	501	14	610	15	19,62	592	519	91
11	Mei 2023	599	20	685	14	15,29	666	601	84
12	Juni 2023	603	17	735	14	16,81	717	647	88
13	Juli 2023	629	20	648	13	15,34	618	562	86
14	Agustus 2023	654	18	659	13	13,47	641	572	87
15	September 2023	619	20	670	14	15,81	643	582	88
16	Oktober 2023	601	22	501	16	15,49	481	436	65
17	November 2023	566	22	599	15	15,29	578	529	70

18	Desember 2023	528	18	603	15	15,40	579	528	75
----	---------------	-----	----	-----	----	-------	-----	-----	----

Setelah dilakukan uji variabel melalui uji asumsi klasik dan uji hipotesis pada 7 variabel bebas tersebut, variabel bebas belum memenuhi syarat uji asumsi klasik terutama dalam Uji Multikolinieritas, ditunjukkan pada Tabel 3 berikut.

Tabel 3. Uji Multikolinieritas

		Coefficients ^a						
Model		Unstandardized		Standardized	t	Sig.	Collinearity Statistics	
		Coefficients		Coefficients			Tolerance	VIF
		B	Std. Error	Beta				
1	(Constant)	760,011	227,568		3,340	,007		
	X ₁	12,271	3,995	,351	3,072	,011	,586	1,706
	X ₃	-23,352	10,170	-,379	-2,296	,042	,281	3,562
	X ₄	-11,135	6,004	-,272	-1,854	,091	,354	2,822
	X ₅	,830	1,948	,873	,426	,678	,002	550,363
	X ₇	-,815	1,919	-,767	-,424	,679	,002	427,848
	X ₈	,691	2,116	,113	,327	,750	,064	15,621

a. *Dependent Variable: Y*

Warna merah pada Table 3 menunjukkan nilai variabel bebas tidak memenuhi syarat uji multikolinieritas (Nilai *Tolerance* > 0,10 atau Nilai *Variance Inflation Factor* (VIF) < 10). Hal ini berarti 7 variabel bebas tersebut masih ada keterkaitan antar variabel, bahkan X₂ tidak memiliki nilai multikolinieritas yang berarti terjadi multikolinieritas sempurna.

Permasalahan ini dapat diatasi dengan menghilangkan /melepas salah satu variabel yang berkaitan tersebut. Ini bisa dilakukan dengan melakukan perulangan uji asumsi klasik dengan menghilangkan salah satu variabel secara acak sampai dihasilkan variabel-variabel bebas yang lulus uji asumsi klasik, untuk selanjutnya dipilih variabel terbaik dengan koefisien determinasi tertinggi dan nilai *error* regresi terkecil melalui uji hipotesis. Nilai *error* regresi digunakan untuk mengukur tingkat penyimpangan masing-masing koefisien regresi. Semakin kecil nilai kesalahan baku, semakin signifikan peran variabel tersebut dalam model, dan sebaliknya (Ir. Tohap Parulian, 2020). Hasilnya terlihat pada Tabel 4 berikut.

Tabel 4. Hasil Perulangan Uji Variabel

Uji	VT	V B	Nor	Mul	Het	Aut	Uji-T	Uji-R	Det	Error Regresi	L/TL
			> 0,05	T > 0,10	VIF < 10	> 0,05	> 0,05	< 0,05	< 0,05		
1	Y	X ₁	0,200	0,586	1,706	0,163		0,011	91,6 %	29,512	TL
		X ₂		0,000	0,000	0,000		0,000			
		X ₃		0,281	3,562	0,683		0,042			
		X ₄		0,354	2,822	0,789	0,466	0,091			
		X ₅		0,002	550,363	0,841		0,678			
		X ₇		0,002	427,848	0,837		0,679			
		X ₈		0,064	15,621	0,780		0,750			
2	Y	X ₁	0,200	0,586	1,706	0,163		0,011	91,6 %	29,512	TL
		X ₃		0,281	3,562	0,683		0,042			
		X ₄		0,354	2,822	0,789	0,466	0,091			
		X ₅		0,002	550,363	0,841		0,678			
		X ₇		0,002	427,848	0,837		0,679			
		X ₈		0,064	15,621	0,780		0,750			
3	Y	X ₁	0,200	0,665	1,503	0,137		0,007	91,5 %	28,488	TL
		X ₂		0,000	0,000	0,000	0,466	0,000			
		X ₃		0,284	3,520	0,982		0,031			
		X ₄		0,370	2,705	0,903		0,062			

4	Y	0,200	X7	0,537	1,864	0,948	0,993	0,466	0,000	91,6 %	29,512	TL
			X8	0,398	2,510	0,129	0,089					
			X1	0,586	1,706	0,163	0,011					
			X2	0,002	537,434	0,837	0,679					
			X3	0,281	3,562	0,683	0,042					
			X4	0,354	2,822	0,789	0,091					
			X5	0,002	550,363	0,841	0,678					
			X8	0,347	2,883	0,290	0,126					
5	Y	0,200	X1	0,586	1,706	0,163	0,011	0,466	0,000	91,6 %	29,512	TL
			X2	0,002	653,080	0,780	0,750					
			X3	0,281	3,562	0,683	0,042					
			X4	0,354	2,822	0,789	0,091					
			X5	0,002	550,363	0,841	0,678					
			X7	0,010	95,949	0,290	0,126					
			X1	0,665	1,503	0,137	0,007					
			X3	0,284	3,520	0,982	0,031					
6	Y	0,200	X4	0,370	2,705	0,903	0,062	0,466	0,000	91,5 %	28,488	L
			X7	0,537	1,864	0,948	0,993					
			X8	0,398	2,510	0,129	0,089					
			X1	0,662	1,510	0,138	0,007					
			X3	0,284	3,526	0,968	0,031					
			X4	0,368	2,715	0,910	0,063					
			X5	0,417	2,397	0,884	0,970					
			X8	0,347	2,883	0,166	0,113					
7	Y	0,200	X1	0,615	1,627	0,145	0,006	0,466	0,000	91,5 %	28,486	L
			X3	0,283	3,530	0,431	0,032					
			X4	0,372	2,691	0,573	0,082					
			X5	0,011	88,429	0,388	0,085					
			X7	0,013	78,951	0,414	0,107					
			X1	0,665	1,503	0,137	0,007					
			X2	0,427	2,341	0,948	0,993					
			X3	0,284	3,520	0,982	0,031					
8	Y	0,200	X4	0,370	2,705	0,903	0,062	0,466	0,000	91,5 %	28,488	L
			X8	0,347	2,881	0,157	0,110					
			X1	0,665	1,503	0,137	0,007					
			X2	0,010	104,933	0,129	0,089					
			X3	0,284	3,520	0,982	0,031					
			X4	0,370	2,705	0,903	0,062					
			X7	0,010	95,886	0,157	0,110					
			X1	0,614	1,628	0,264	0,023					
9	Y	0,200	X2	0,002	537,385	0,880	0,709	0,808	0,000	89,5 %	31,585	TL
			X3	0,347	2,885	0,353	0,009					
			X4	0,357	2,800	0,585	0,135					
			X5	0,002	550,004	0,827	0,669					

Hasil Uji Variabel menunjukkan hasil terbaik pada percobaan ke 7 (warna hijau) dengan menghilangkan 2 variabel bebas yaitu X2 dan X7 dengan koefisien determinasi 91,5% dan nilai *error* regresi sebesar 28,486, artinya 5 variabel bebas tersebut secara simultan memiliki pengaruh 91,5% terhadap variabel terikat, dan 8,5% dipengaruhi oleh variabel lain dengan tingkat kesalahan regresi terkecil, sehingga berdasarkan uji variabel diperoleh variabel bebas yang tepat adalah variabel bebas X1 (Jumlah Hari Kerja), X3 (Usia Kendaraan 6 bulan sebelumnya), X4 (Jarak Tempuh/Alamat 6 bulan sebelumnya), X5 (Jenis Kendaraan Mobil Barang 6 bulan sebelumnya) dan X8 (Sifat Kendaraan Umum 6 bulan sebelumnya) dengan persamaan ditunjukkan pada Tabel 5 berikut.

Tabel 5. Persamaan Regresi Linier Berganda

Model	Unstandardized Coefficients		
		B	Std. Error
1	(Constant)	773,884	217,379
	X ₁	11,695	3,627

X ₃	-23,788	9,766
X ₄	-11,633	5,684
X ₅	,005	,124
X ₈	1,502	,877

$$Y=773,884+11,695X_1-23,788X_3-11,633X_4+0,005X_5+1,502X_8.....(3)$$

Berdasarkan persamaan Regresi Linier Berganda (3), dapat dilakukan peramalan terhadap jumlah kendaraan diuji baik periode bulan Januari 2024 sampai dengan bulan Juni 2024 dengan data variabel seperti ditunjukkan pada Tabel 6 berikut.

Tabel 6. Data Variabel Bebas Untuk Dilakukan Peramalan

No.	Bulan	Y	X ₁	X ₃	X ₄	X ₅	X ₈
1	Januari 2024	607	21	15	14,41	605	68
2	Februari 2024	460	15	16	14,47	627	91
3	Maret 2024	611	17	14	13,88	607	76
4	April 2024	559	16	14	14,06	574	73
5	Mei 2024	606	18	15	15,16	542	77
6	Juni 2024	570	22	15	15,96	510	60

Dengan menggunakan aplikasi Ms. Exel, maka perhitungan peramalan dan uji akurasi diperoleh hasil seperti ditunjukkan pada Tabel 7 berikut.

Tabel 7. Peramalan dan Uji Akurasi

No	Bulan	Y	X ₁	X ₃	X ₄	X ₅	X ₈	Prediksi	Error	ABS (Error)	ABS(Error)/Y
1	Januari 2024	607	21	15	14,41	605	68	600	7	7	0,011221631
2	Februari 2024	460	15	16	14,47	627	91	540	- 80	80	0,174322804
3	Maret 2024	611	17	14	13,88	607	76	595	16	16	0,025551620
4	April 2024	559	16	14	14,06	574	73	577	- 18	18	0,032071592
5	Mei 2024	606	18	15	15,16	542	77	570	36	36	0,060096172
6	Juni 2024	570	22	15	15,96	510	60	581	- 11	11	0,019932140
Jumlah =									- 51	168	0,32319596

Diperoleh hasil peramalan sebagai berikut :

1. Bulan Januari 2024 = 600 unit kendaraan
2. Bulan Februari 2024 = 540 unit kendaraan
3. Bulan Maret 2024 = 595 unit kendaraan
4. Bulan April 2024 = 577 unit kendaraan
5. Bulan Mei 2024 = 570 unit kendaraan
6. Bulan Juni 2024 = 581 unit kendaraan

Sehingga dengan menggunakan uji akurasi MAPE diperoleh nilai akurasi sebagai berikut :

$$MAPE = \left(\frac{1}{N} \sum_{t=1}^n \frac{|y_t - \hat{y}_t|}{y_t} \right) \times 100\%$$

$$MAPE = \left(\frac{1}{6} 0,32319596 \right) \times 100\% = 5,39\%$$

KESIMPULAN

Peramalan jumlah kendaraan diuji baik merupakan solusi dari permasalahan penyediaan bahan baku PKB Dishub Buleleng. Variabel yang tepat yang mempengaruhi jumlah kendaraan diuji baik terdiri dari 5 variabel bebas yaitu Jumlah hari kerja dibulan bersangkutan, Usia kendaraan dari kendaraan diuji baik 6 bulan sebelumnya, Jarak tempuh (alamat pemilik) dari kendaraan diuji baik 6 bulan sebelumnya, Jenis kendaraan mobil barang dari kendaraan diuji baik 6 bulan sebelumnya, dan Sifat kendaraan umum dari kendaraan diuji baik 6 bulan sebelumnya. Dengan persamaan Regresi Linier Berganda $Y = 773,884 + 11,695X_1 - 23,788X_2 - 11,633X_3 + 0,005X_4 + 1,502X_5$ dihasilkan hasil peramalan jumlah kendaraan diuji baik dari bulan Januari 2024 sampai dengan bulan Juni 2024 secara berturut-turut adalah 600, 540, 595, 577, 570 dan 581 kemudian dilakukan uji akurasi dan menghasilkan Nilai *Mean Absolute Percentage Error* (MAPE) sebesar 5,39 dengan kategori Sangat Bagus. Karena dalam penelitian ini, untuk variabel Jarak tempuh (alamat pemilik) masih menggunakan sumber data sekunder berbasis desa/kelurahan, sehingga untuk penelitian selanjutnya, dapat lebih meningkatkan hasil penelitian dengan lebih mengakuratkan variabel bebas Jarak tempuh (alamat pemilik) dengan mengambil koordinat titik lokasi secara lebih akurat.

DAFTAR PUSTAKA

- Ahmad, N., Ghadi, Y., Adnan, M., & Ali, M. (2022). Load forecasting techniques for power system: Research challenges and survey. *IEEE Access*, 10, 71054-71090. <https://doi.org/10.1109/ACCESS.2022.3187839>
- Algoritma. (2022, February). *Pentingnya Data Cleaning Dalam Data Science*. Algoritma. <https://algoritma.blog/data-cleaning-adalah/>
- Allo, G. (2023). Perbandingan Metode Feature Selection Filter Method Dan Backward Method Pada Dataset Dermatology. *Cenderawasih : Journal of Statistics and Data Science*, 2(1), 32-36.
- Azizah, S. R., Herteno, R., Farmadi, A., Kartini, D., & Budiman, I. (2023). Kombinasi Seleksi Fitur Berbasis Filter dan Wrapper Menggunakan Naive Bayes pada Klasifikasi Penyakit Jantung. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, 10(6), 1361-1368. <https://doi.org/10.25126/jtiik.1067467>
- Hanafi, H. (2023). *Data Cleaning dalam Big Data : Review*. <https://www.researchgate.net/publication/376758940>
- Mardiatmoko, G.-. (2020). The Importance of the Classical Assumption Test in Multiple Linear Regression Analysis (A Case Study of the Preparation of the Allometric Equation of Young Walnuts). *BAREKENG: Jurnal Ilmu Matematika Dan Terapan*, 14(3), 333-342. <https://doi.org/10.30598/barekengvol14iss3pp333-342>
- Marta, T., Mulyono, H., & Irsyadunas. (2023). Analisis Penerimaan Siswa Terhadap Penggunaan Google Classroom Dengan Metode Technology Acceptance Model (TAM). *Decode: Jurnal Pendidikan Teknologi Informasi*, 3(1), 30-37. <https://doi.org/10.51454/decode.v3i1.71>
- Nur, M., Rizki, E. N., Karim, A. A., & Sari, R. K. (2024). Peramalan Jumlah Penumpang Domestik Pada Bandar Udara Sultan Syarif Kasim II Dengan Menggunakan Metode Winter's Exponential Smoothing. *Jurnal Teknologi dan Manajemen Industri Terapan*, 3(I), 57-66. <https://doi.org/10.55826/tmit.v3ii.302>
- Purba, D. S., Tarigan, W. J., Sinaga, M., & Tarigan, V. (2021). Pelatihan Penggunaan Software SPSS Dalam Pengolahan Regresi Linear Berganda Untuk Mahasiswa Fakultas Ekonomi Universitas Simalungun Di Masa Pandemi Covid 19. *Jurnal Karya Abadi*, 5(2), 202-208.

- Rahma, V. A., & Sihombing, T. (2023). Implementasi Kebijakan Pengujian Kendaraan Bermotor Jenis Angkutan Kota Untuk Meningkatkan Keselamatan Penumpang di Kota Medan. *Journal Of Science and Social Research*, 6(1). <https://doi.org/10.54314/jssr.v6i1.1191>
- Ramadhani, R. A., Maharani, M. D., Dewita, N. A. P., Nanlohy, S. A. R., & Maesaroh, S. S. (2023). Pemanfaatan Aplikasi SPSS Dalam Analisis Pengaruh Biaya Pemasaran Dan Gaji Pada Pendapatan Perusahaan Telekomunikasi. *JIKA (Jurnal Informatika)*, 7(3), 271. <https://doi.org/10.31000/jika.v7i3.8328>
- Rismanitanti, A., Mawarni, R., Rahmatullah, S., Efendi, D. M., & Nurbaiti, S. (2022). Perbandingan Pengolahan Data Prediksi Persediaan Gas Lpg 3kg Menggunakan Regresi Linier Berganda Dan K-Means. *Jurnal Informatika*, 22(2), 183-195.
- Rohaeti, E., Andriyati, A., Sumarsa, A., & Widyastiti, M. (2022). Peramalan Dan Penjadwalan Ketersediaan Bahan Baku Untuk Mengoptimalkan Perencanaan Produksi Usaha Kecil Menengah Di Bogor. *Jurnal PkM Pengabdian Kepada Masyarakat*, 5(1). <https://doi.org/10.30998/jurnalpkm.v5i1.7723>
- Ruspendi, Rusmalah, & Syahreen, N. (2022). *Teknik Peramalan* (1st ed.). Unpam Press.
- Sabtohadhi, J., Fatimah, F., Mei Saputri, D., Agus Tina, L., Novitasari, H., Rofi, I., Fadlilah, F., Nabila Imsaki Aulia, S., Indah Romadloni, R., & Resti Universitas Muhammadiyah Kalimantan Timur, L. (2024). Classic Assumption Data Analysis Training Using the SPSS Application. In *Community Engagement & Emergence Journal* (Vol. 5, Issue 1).
- Susilawati, R., & Sunendiari, S. (2022). Peramalan Jumlah Penumpang Kereta Api Menggunakan Metode Arima dan Grey System Theory. *Jurnal Riset Statistika*, 2(1), 1-12. <https://doi.org/10.29313/jrs.vi.603>
- Tania, M., Dalimunthe, R. A., & Azmi, S. R. M. (2024). Penerapan Metode Single Moving Average Untuk Memprediksi Jumlah Pertumbuhan Penduduk. *Decode: Jurnal Pendidikan Teknologi Informasi*, 4(2), 452-465. <https://doi.org/10.51454/decode.v4i2.490>
- Trunfio, T. A., Scala, A., Giglio, C., Rossi, G., Borrelli, A., Romano, M., & Improta, G. (2022). Multiple regression model to analyze the total LOS for patients undergoing laparoscopic appendectomy. *BMC Medical Informatics and Decision Making*, 22(1), 141. <https://doi.org/10.1186/s12911-022-01884-9>
- Urfah. (2022). *Mengenal Feature Selection dalam Machine Learning*. Datalearns247. <https://www.datalearns247.com/mengenal-feature-selection-dalam-machine-learning-69>
- Utami, A. A., Hernawan, D., & Purnamasari, I. (2020). Analysis Of Service Quality In Testing Goods Transportation Type Motor Vehicle. *Jurnal Governansi*, 6(2). <https://doi.org/10.30997/jgs.v6i2.3042>