

Prediksi Jumlah Kendaraan Bermotor Menggunakan Machine Learning

Ahmad Darmawan Sidik ^{1*}, Alek Ansawarman ²
Universitas Tama Jagakarsa

ABSTRACT: The number of motorized vehicles every year is always increasing .This increase needs to be considered because motorized vehicles require road facilities. In building road facilities, careful planning is needed in order to accommodate the growth of motorized vehicles in the future. By making predictions, the number of vehicles in the future can be known. In this study, predictions were made using machine learning with a linear regression model. To make it easier to see the prediction results, a website was created using flask in python. From the test results, it is found that the model made has an R2 accuracy value of 0.98. These results indicate that the regression model made is very good.

Keywords: The number of motorized vehicles, regression model, road facilities

ABSTRAK: Jumlah kendaraan bermotor setiap tahun selalu meningkat. Peningkatan ini perlu diperhatikan karena kendaraan bermotor membutuhkan sarana jalan. Dalam membangun sarana jalan diperlukan perencanaan yang matang agar dapat menampung pertumbuhan kendaraan bermotor di masa yang akan datang. Dengan melakukan prediksi maka jumlah kendaraan di masa yang akan datang dapat diketahui. Dalam penelitian ini prediksi dilakukan menggunakan machine learning dengan model regresi linear. Untuk memudahkan melihat hasil prediksinya dibuat website menggunakan flask pada python. Dari hasil pengujian didapat bahwa model yang dibuat memiliki nilai akurasi R2 sebesar 0,98. Hasil ini menunjukkan bahwa model regresi yang dibuat sangat baik.

Kata Kunci : Jumlah Kendaraan Bermotor, Model Regresi, Fasilitas Jalan

Submitted: 07-07-2022; Revised: 13-07-2022; Accepted:23-07-2022

Corresponding Author: Sidik@gmail.com

PENDAHULUAN

Seiring dengan meningkatnya jumlah penduduk, kebutuhan sarana transportasi juga makin meningkat. Kebutuhan sarana transportasi ini dipenuhi dengan berbagai jenis kendaraan motor. Berdasarkan data BPS jumlah kendaraan bermotor di Indonesia pada tahun 2021 sebanyak 143.797.227. Jumlah tersebut meningkat 5,7% dibanding tahun sebelumnya. Informasi peningkatan jumlah kendaraan bermotor perlu diketahui karena kendaraan bermotor membutuhkan sarana jalan. Peningkatan jumlah kendaraan tanpa dibarengi oleh penambahan sarana jalan akan mengakibatkan timbulnya kemacetan yang parah. Dalam menambah sarana jalan perlu dilakukan perencanaan matang. Diantaranya dengan memprediksi jumlah kendaraan bermotor dimasa mendatang. Dengan mengetahui jumlah kendaraan bermotor di masa mendatang maka dapat diketahui seberapa besar sarana jalan yang dibutuhkan. Disamping itu informasi jumlah kendaraan bermotor di masa mendatang bisa juga dijadikan masukan dalam perencanaan tata kota dan pengembangan wilayah serta regulasi lalu lintas.

TINJAUAN PUSTAKA

2.a. Machine Learning

Machine learning merupakan bidang lain dari ilmu komputer yang merancang sebuah algoritma agar memungkinkan sebuah komputer untuk belajar melalui data sehingga sering dikatakan sebagai learn from data. Jadi machine learning adalah pemrograman komputer yang menggunakan data masa lalu yang digunakan untuk pembelajaran model sehingga mendapatkan performa yang optimal dalam menggali informasi dari suatu kumpulan data[1]. Inti machine learning adalah untuk membuat model yang merefleksikan pola-pola data. Sedangkan menurut Tom M. Mitchell definisi Machine Learning adalah sebuah program komputer yang belajar dari pengalaman dari tugas yang dibebankan dengan kinerjanya yang terukur.

Secara garis besar machine learning dibagi menjadi 3 yaitu Supervised Learning, Unsupervised Learning dan Reinforcement Learning.

- a. Supervised Learning adalah algoritma machine learning yang proses pembelajarannya dibawah pengawasan. Yang termasuk dalam supervised learning diantaranya adalah classification dan regression.
- b. Unsupervised Learning adalah algoritma machine learning yang proses pembelajarannya tanpa pengawasan. Yang termasuk dalam unsupervised learning diantaranya adalah clustering dan dimensionality reduction.
- c. Reinforcement Learning adalah algoritma machine learning yang bisa membuat agent software mesin bekerja dan bekerja secara otomatis untuk

menentukan perilaku ideal sehingga dapat memaksimalkan kinerja algoritma. Yang termasuk dalam Reinforcement Learning diantaranya adalah real-time decision, robot navigation, learning tasks, skills acquisition dan game AI.

2.b Regresi

Regresi adalah metode statistik yang dipakai untuk memperkirakan hubungan antara sebuah variabel terkait dengan satu variabel independen atau lebih. Metode ini juga bisa digunakan untuk menilai kekuatan hubungan antara variabel dengan perkiraan masa depan.

Regresi linear sederhana merupakan suatu model persamaan yang menggambarkan hubungan satu variabel bebas (X) dengan satu variabel tak bebas (Y).

Persamaan regresi linear sederhana secara matematik diekspresikan oleh

$$\hat{Y} = a + bX$$

Dimana

\hat{Y} = garis regresi

a = konstanta (intersep)

b = konstanta regresi (slope)

X = variabel bebas (predictor)

Besarnya konstanta a dan b dapat ditentukan menggunakan persamaan :

$$a = \frac{(\sum Y_i)(\sum X_i^2) - (\sum X_i)(\sum X_i Y_i)}{n \sum X_i^2 - (\sum X_i)^2}$$

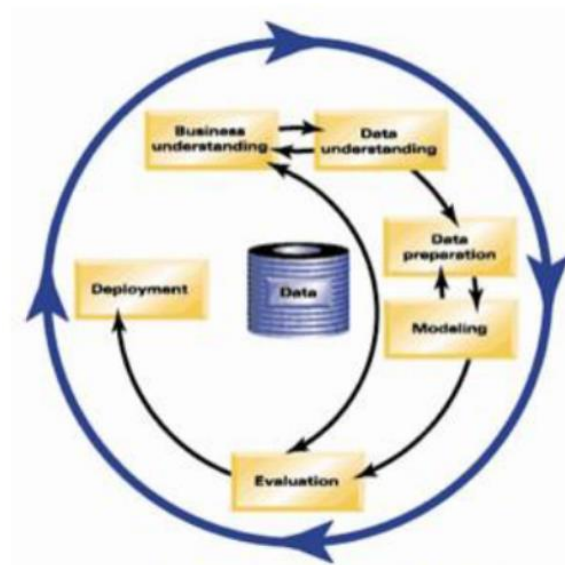
$$b = \frac{n (\sum X_i Y_i) - (\sum X_i)(\sum Y_i)}{n \sum X_i^2 - (\sum X_i)^2}$$

Dimana

n = jumlah data

METODOLOGI

Dalam penelitian ini metode yang digunakan adalah CRISP DM (Cross Industry Standard Processing for Data Mining).



Gambar 1 Standar CRISP-DM

CRISP-DM merupakan standar yang dikembangkan di Eropa yang terdiri atas 6 tahap yaitu

a. Business Understanding

Merupakan tahap dimana permasalahan didefinisikan dengan tepat

b. Data Understanding

Proses dimana mempertemukan antara data yang kita miliki dan data apa yang seharusnya diperlukan

c. Data Preparation

Proses persiapan data untuk diolah menggunakan model yang tepat.

d. Modeling

Pembuatan model sesuai dengan karakteristik data, tujuan teknis dan parameter model.

e. Evaluasi

Menguji model yang dibuat berdasarkan data dan parameter yang relevan.

f. Deployment

Mengembangkan model yang sudah dibuat dengan mempertimbangkan diantaranya kemudahan penggunaan model dan pemeliharannya di masa mendatang.

PEMBAHASAN

4.a Business Understanding

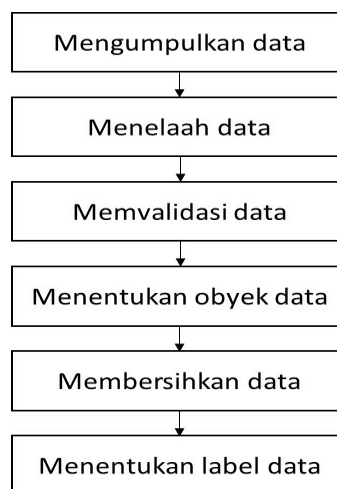
Dalam penelitian ini permasalahan terkait dengan meningkatnya jumlah kendaraan bermotor yang berdampak pada padatnya volume kendaraan bermotor di jalan yang pada akhirnya bisa berakibat kemacetan. Dengan demikian informasi jumlah kendaraan bermotor di masa mendatang menjadi penting sekali karena terkait dengan perencanaan sarana jalan, pengembangan dan penataan kota serta regulasi lalu lintas.

4.b Data Understanding

Output dari penelitian ini adalah prediksi jumlah kendaraan bermotor di masa mendatang. Dataset yang dibutuhkan adalah jumlah kendaraan bermotor pada tahun-tahun di masa lalu secara lengkap dan akurat.

4.c Data Preparation

Dalam penelitian ini sumber data diambil dari Badan Pusat Statistik. Data yang berhasil dikumpulkan tersebut selanjutnya diproses dalam beberapa tahap agar dapat diperoleh data yang bersih dan valid.

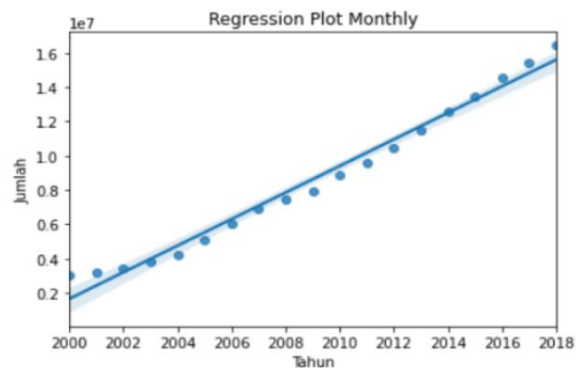


Gambar 2. Tahapan pengolahan data

4.d Modeling

Pemodelan dilakukan menggunakan Jupyter Notebook dengan bahasa pemrograman Python.

Dalam tahap modeling ini dilakukan visualisasi atas data yang sudah disiapkan.



Gambar 3 Visualisasi data

Dalam penelitian ini hasil yang diinginkan adalah nilai prediksi. Menentukan prediksi pada Machine Learning termasuk dalam Supervised Learning dengan model regresi.

Model regresi adalah hasil dari proses penentuan parameter regresi yang dapat dengan akurat memprediksi variabel tak bebas berdasar variabel bebas. Dalam penelitian ini variabel bebasnya adalah tahun sedangkan variabel tak bebasnya adalah jumlah kendaraan bermotor.

Model yang digunakan adalah jenis model Regresi Linear. Kelebihan model regresi linear adalah tidak perlu tuning parameter, proses komputasi nya sederhana dan mudah dipahami dan diinterpretasikan.

Tabel 1. Hasil prediksi jumlah kendaraan mobil hingga tahun 2030.

Tahun	Jumlah kendaraan mobil
2022	18714651
2023	19490565
2024	20266480
2025	21042395
2026	21818310
2027	22594225
2028	23370140
2029	24146054
2030	24921969

Model Regression Linear yang sudah dibuat di Jupyter Notebook selanjutnya disimpan dalam bentuk file.

4.e Evaluation

Proses evaluasi merupakan hal yang penting dalam mengembangkan model machine learning. Pada model regresi pengujian dapat dilakukan dengan metode R^2 . R^2 merupakan suatu nilai yang memperlihatkan seberapa besar variabel independen mempengaruhi variabel dependen. Semakin mendekati angka satu, maka hasil dari model regresi akan semakin baik.

Rumus dari R^2 adalah

$$R^2 = 1 - \frac{SS\ Error}{SS\ Total} = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2}$$

Dimana

y_i = observasi respon ke-i

\bar{y} = rata-rata

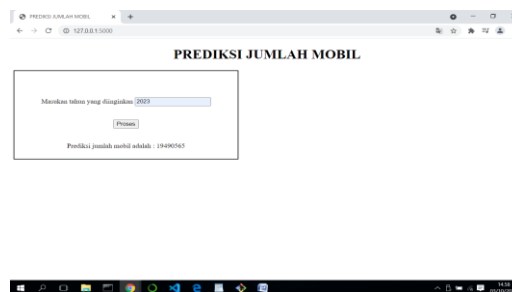
\hat{y}_i = prediksi respon ke-i

Dengan menggunakan bahasa pemrograman Python didapat nilai R^2 dari model yang dibangun sebesar 0,98. Hasil ini menunjukkan bahwa model regresi memiliki akurasi yang baik.

4.f Deployment

Setelah dilakukan pengujian selanjutnya adalah tahap pengembangan model. Pada tahap dilakukan pembuatan website untuk memudahkan melihat hasil prediksi sesuai tahun yang diinginkan.

Pembuatan website menggunakan framework Flask pada Python. Flask adalah aplikasi dalam bahasa pemrograman Python yang menggunakan toolkit Werkzeug dan template Jinja2. Werkzeug berperan dalam menyediakan routing, debugging dan Web Server Interface Gateway, sedangkan Jinja2 sebagai template engine. Untuk proses pembuatannya dibutuhkan juga software pendukung Anaconda dan Heroku. Di Anaconda terdapat kumpulan aplikasi, library dan tool yang digunakan untuk pengembangan program. Sementara Heroku merupakan layanan PaaS yang menyediakan pelayanan web untuk menempatkan aplikasi yang sudah dikembangkan agar semua orang dapat mengaksesnya. Dengan dukungan Git aplikasi yang telah dibuat bisa ditempatkan di layanan Heroku sehingga bisa diakses dimana saja melalui internet.



Gambar 4 Tampilan aplikasi di Heroku

KESIMPULAN DAN REKOMENDASI

- a. Dari hasil pengujian yang dilakukan pada proses evaluasi dapat diketahui bahwa model machine learning yang dibuat memiliki nilai akurasi R^2 sebesar 0,98. Nilai ini menunjukkan bahwa model memiliki akurasi yang baik.
- b. Hasil prediksi di tahun-tahun yang akan datang menunjukkan jumlah kendaraan bermotor makin meningkat. Perlu perhatian yang serius dari pihak yang berwenang untuk mengantisipasi hal tersebut agar tidak terjadi kemacetan yang parah.

DAFTAR PUSTAKA

Alpaydin, E (2010), Introduction to Machine Learning, London, MIT Press.

Digital Talent.Scholarship, Modul Pelatihan AI Data Science, Kementrain Komunikasi dan Informatika, 2021

Dataindonesia.id,<https://dataindonesia.id/sektor-riil/detail/jumlah-kendaraan-bermotor-di-8-provinsi-ini-paling-sedikit>, 2021

Badan Pusat Statistik, 2021

Hakim, Nafisah, Implementasi Machine Learning pada system prediksi kejadian dan lokasi patah rel kereta api di Indonesia, Jurnal Sistem Cerdas 2020. Vol 03

SKKNI Bidang Keahlian Artificial Intelligence Subbidang Data Science. Kementerian Departemen Tenaga Kerja. 2020.

Yuliara, Made, Regresi Linear Sederhana, Universitas Udayana.