

Rangkuman Perkuliahan: Implementasi Linear Regression Menggunakan Pendekatan Analitik dan Numerik Berbasis Python NumPy

Muhammad Labib¹, Muhammad Shadaam Haidar Yuwono²

^{1,2} Program Studi Teknik Informatika, STMIK Tazkia Bogor
E-mail : 241552010005.shadaam@student.stmik.tazkia.ac.id

Abstrak

Linear Regression merupakan salah satu metode dasar dalam machine learning dan statistika yang digunakan untuk memodelkan hubungan linier antara variabel independen dan variabel dependen. Penelitian ini bertujuan untuk menjelaskan konsep dasar linear regression, metode perhitungan parameter secara analitik, serta implementasi numerik menggunakan bahasa pemrograman Python dengan library NumPy. Pendekatan analitik digunakan untuk menghitung nilai gradien dan intercept secara matematis, sedangkan pendekatan numerik digunakan untuk mempermudah komputasi pada dataset yang lebih besar. Hasil pembahasan menunjukkan bahwa penggunaan NumPy mampu menyederhanakan proses perhitungan, meningkatkan efisiensi komputasi, serta memudahkan visualisasi hasil regresi. Penelitian ini diharapkan dapat menjadi referensi awal bagi mahasiswa dalam memahami regresi linier baik secara teoritis maupun implementatif.

Kata kunci: Linear Regression, Machine Learning, NumPy, Analitik, Numerik

1. Pendahuluan

Pemanfaatan data dalam berbagai bidang ilmu pengetahuan semakin meningkat seiring dengan perkembangan teknologi informasi. Salah satu metode analisis data yang paling dasar dan banyak digunakan adalah Linear Regression. Metode ini bertujuan untuk memodelkan hubungan linier antara variabel independen dan variabel dependen. Dalam konteks machine learning, linear regression sering digunakan sebagai dasar untuk memahami konsep optimasi, error, dan prediksi sebelum mempelajari algoritma yang lebih kompleks.

2. Landasan Teori

2.1 Linear Regression

Linear regression dengan satu variabel independen dapat dinyatakan dalam bentuk persamaan:

$$y = mx + c$$

Dalam notasi machine learning, persamaan tersebut dapat dituliskan sebagai:

$$f(x) = w_1x + w_0$$

di mana w_1 merupakan bobot (gradien) dan w_0 merupakan bias atau intercept.

2.2 Perhitungan Parameter Secara Analitik

Nilai gradien dan intercept pada regresi linier satu variabel dapat dihitung menggunakan rumus statistik sebagai berikut:

$$m = \frac{n \sum(xy) - \sum x \sum y}{n \sum(x^2) - (\sum x)^2}$$

$$c = \bar{y} - m\bar{x}$$

Pendekatan ini efektif digunakan untuk dataset berukuran kecil dan membantu memahami dasar matematis dari linear regression.

2.3 Pendekatan Numerik

Pendekatan numerik digunakan untuk dataset yang lebih besar dengan memanfaatkan proses iteratif. Salah satu metode yang umum digunakan adalah Gradient Descent, yang bekerja dengan meminimalkan nilai error secara bertahap melalui pembaruan parameter model.

3. Implementasi Menggunakan Python NumPy

NumPy merupakan library Python yang mendukung operasi numerik berbasis array dan matriks. Dengan menggunakan NumPy, perhitungan nilai rata-rata, perkalian vektor, serta perhitungan parameter regresi dapat dilakukan secara efisien tanpa menggunakan perulangan eksplisit. Dataset disimpan dalam format CSV dan diproses menjadi array dua dimensi untuk memisahkan variabel input dan output.

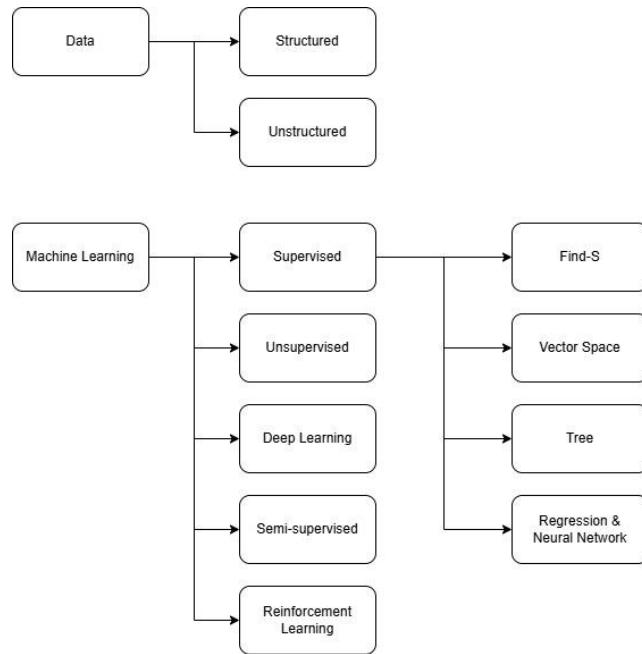
4. Hasil dan Pembahasan

Hasil implementasi menunjukkan bahwa garis regresi yang dihasilkan mampu merepresentasikan pola hubungan linier pada data. Meskipun tidak seluruh titik data berada tepat pada garis regresi, model mampu memberikan pendekatan terbaik berdasarkan minimisasi error. Penggunaan NumPy secara signifikan menyederhanakan kode program dan meningkatkan efisiensi komputasi dibandingkan perhitungan manual.

5. Kesimpulan

Linear regression merupakan metode fundamental dalam analisis data dan machine learning. Pendekatan analitik memberikan pemahaman matematis yang kuat, sedangkan pendekatan numerik memberikan fleksibilitas dalam menangani dataset berukuran besar. Implementasi menggunakan Python NumPy terbukti mempermudah proses komputasi dan sangat sesuai digunakan sebagai media pembelajaran maupun pengembangan aplikasi berbasis data.

6. Mindmap



Daftar Pustaka

- [1] Hastie, T., Tibshirani, R., and Friedman, J., *The Elements of Statistical Learning*, Springer, 2017.
- [2] McKinney, W., *Python for Data Analysis*, O'Reilly Media, 2022.