

Analisis Chapter 4: Training Models

Chapter 4 membahas model linear dan teknik pelatihan model seperti:

- Linear Regression (Normal Equation, Sklearn, SVD, Pseudoinverse)
- Gradient Descent (Batch, Stochastic, Mini-batch)
- Polynomial Regression
- Learning Curves untuk evaluasi overfitting/underfitting
- Regularization: Ridge, Lasso, ElasticNet
- Logistic Regression dan Softmax untuk klasifikasi

Berikut ini analisis dan visualisasi dari setiap bagian eksperimen.

Linear Regression

Model linear digunakan untuk meregresi data bentuk linier. Parameter dihitung melalui:

- Normal Equation
- Scikit-Learn (LinearRegression)
- Least Squares (SVD)
- Pseudoinverse (Moore-Penrose)

Semua metode memberikan hasil prediksi dan parameter yang mirip, cocok untuk dataset kecil.

Gradient Descent

Tiga jenis Gradient Descent diuji:

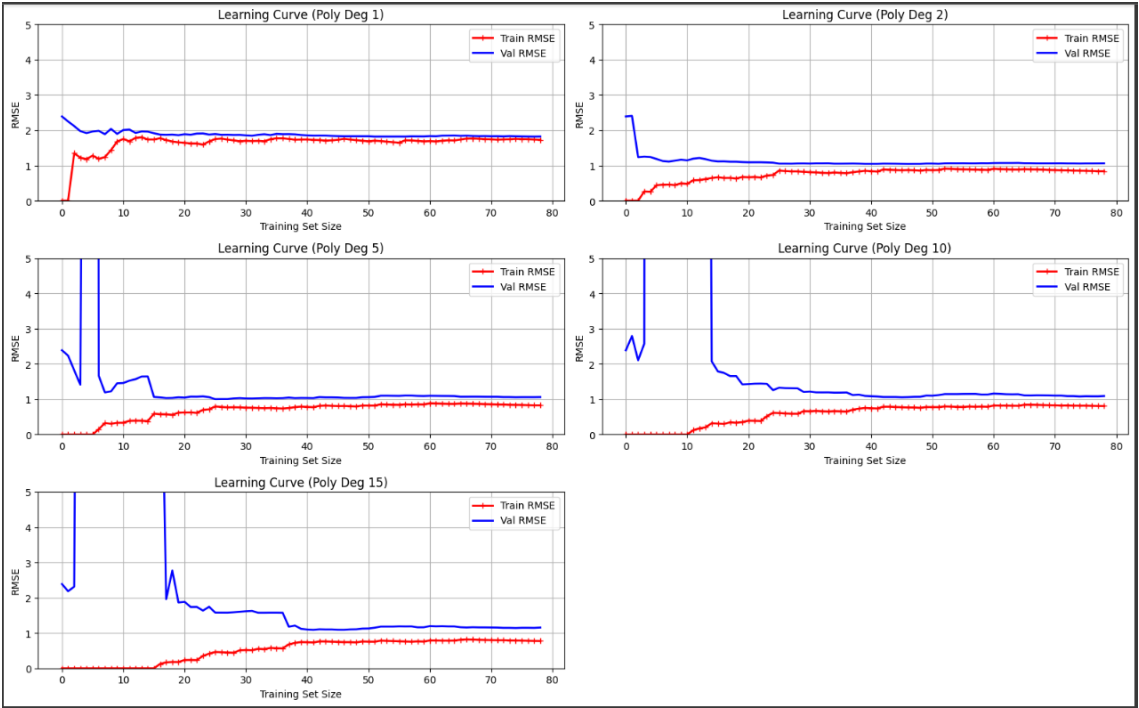
- Batch GD: stabil tapi lambat
- Stochastic GD: cepat tapi fluktuatif
- Mini-batch GD: seimbang antara kecepatan dan kestabilan

Setiap metode sukses mencapai konvergensi ke parameter optimal.

Polynomial Regression

Untuk menangani data non-linear, digunakan PolynomialFeatures (degree=2). Model berhasil menangkap kurva kuadratik dengan baik.

Berikut hasil visualisasi:

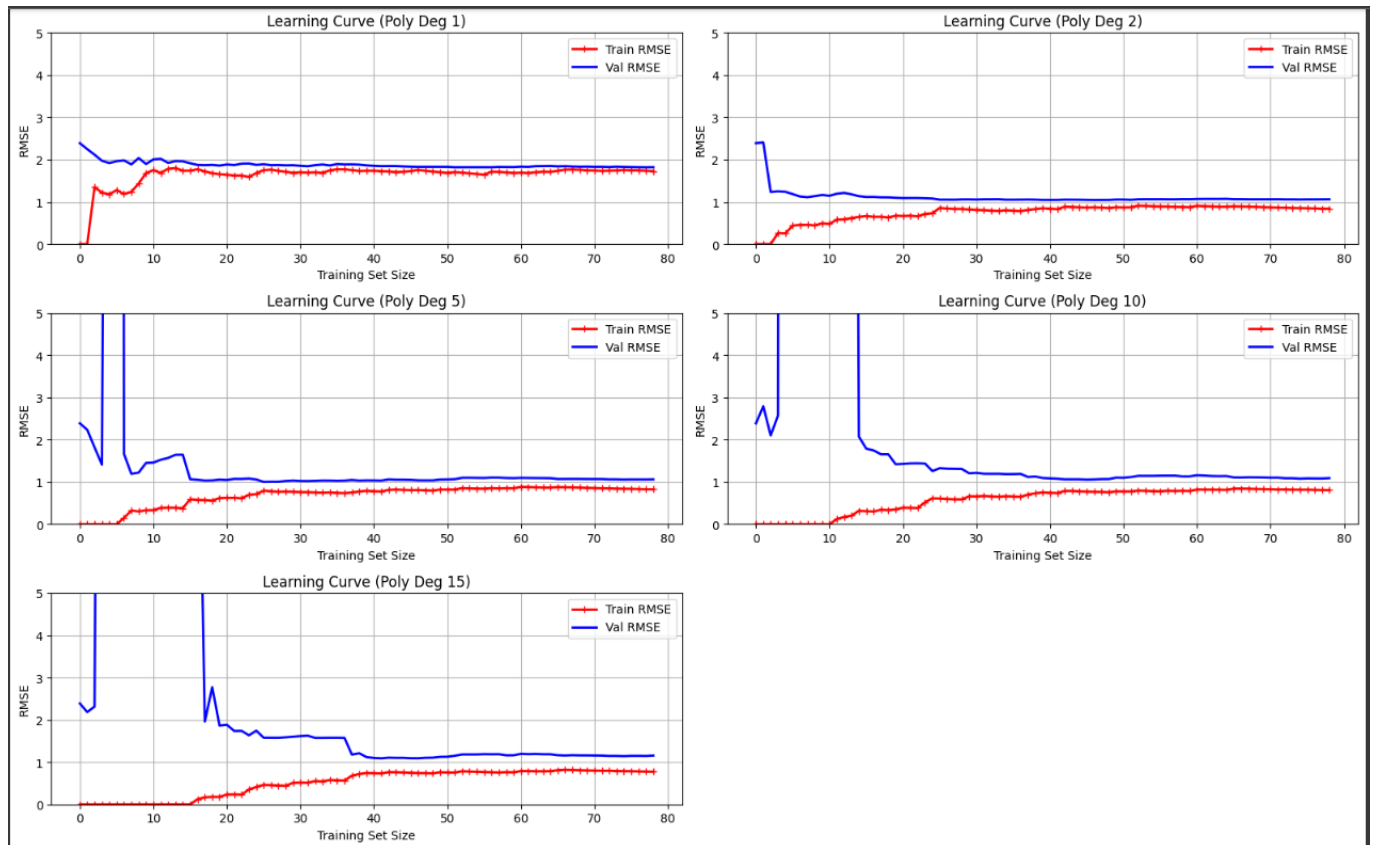


Learning Curves

Analisis pembelajaran dilakukan dengan berbagai derajat polinomial:

- Degree 1: underfitting
- Degree 2-5: good fit
- Degree 10+: overfitting

Visualisasi menunjukkan bahwa model kompleks cenderung overfit terhadap data training.



Regularization (Ridge, Lasso, ElasticNet)

Regularisasi diterapkan pada Polynomial Regression (degree=10):

- Ridge (L2): mengecilkan koefisien, tetap mempertahankan semua fitur
- Lasso (L1): menyusutkan dan menghapus fitur (sparse model)
- ElasticNet: gabungan L1 dan L2

Berguna untuk mengurangi overfitting.

Logistic & Softmax Regression

Model klasifikasi biner:

- Logistic Regression pada fitur petal width untuk memisahkan Virginica
- Visualisasi menunjukkan transisi probabilitas antara kelas

Output Prediksi:

- Prediction: Virginica
- Probabilities: [6.2e-07, 5.74e-02, 94.26%]

Softmax Regression juga digunakan untuk multi-kelas klasifikasi.