

LAPORAN PRAKTIKUM MACHINE LEARNING

Nama	Hidaya
NIM	0030

1. Kode & Hasil Eksperimen

Eksperimen ini menggunakan dataset **Iris** dan model **K-Nearest Neighbors (KNN)** dengan $k=1$. Model diuji menggunakan metode **train_test_split** dan **cross-validation** untuk mengukur akurasi. Hasilnya menunjukkan: Training penuh (fit + predict seluruh data): akurasi **1.0** Split 50%-50%: akurasi **0.9066** Cross Validation (cv=5): akurasi rata-rata sekitar **0.96** Leave-One-Out: akurasi rata-rata **0.96** Selain itu dilakukan juga eksperimen **Polynomial Regression** menggunakan pipeline scikit-learn dengan derajat polinomial bervariasi (degree 1–20). Proses ini melibatkan **validation curve**, **learning curve**, serta **GridSearchCV** untuk menemukan kombinasi terbaik parameter derajat polinomial dan intercept.

2. Grafik dan Interpretasi Hasil

Grafik hasil eksperimen menunjukkan bahwa: Untuk KNN, performa tinggi dan stabil karena dataset Iris relatif sederhana dan terpisah dengan baik. Pada Polynomial Regression, grafik validation curve memperlihatkan bahwa semakin tinggi derajat polinomial menyebabkan **overfitting**, yakni nilai akurasi training meningkat tetapi akurasi validasi menurun. Learning curve menunjukkan bahwa model dengan derajat kecil (degree=2) memiliki bias tinggi, sementara degree besar (degree=9) memiliki varians tinggi. Hasil terbaik (dari GridSearchCV) diperoleh pada degree = 4 dengan fit_intercept = False.

3. Kesimpulan

Dari seluruh hasil pengujian dapat disimpulkan bahwa: Model **KNN** bekerja sangat baik pada dataset Iris dengan akurasi tinggi bahkan pada validasi silang. Untuk **Polynomial Regression**, kompleksitas model berpengaruh besar terhadap performa; derajat polinomial terlalu tinggi menyebabkan overfitting. Penggunaan teknik seperti **cross-validation**, **learning curve**, dan **grid search** sangat membantu dalam mengevaluasi dan memilih model terbaik.