UE23CS352A: MACHINE LEARNING

Week 6: Artificial Neural Networks

Name: Gavini Prithvi Ananya

SRN: PES2UG23CS198

Section: C

Date: 16/09/2025

Purpose of the Lab:
The purpose of this lab was to understand the application of polynomial regression in modelling non-linear relationships between features and target variables. The lab aimed to explore how polynomial transformations improve model fitting for datasets that cannot be accurately modelled with simple linear regression.

Tasks Performed:

- Generated a synthetic dataset based on the assigned polynomial function.
- Implemented polynomial regression using different degrees of polynomials.
- Split the dataset into training and testing sets.
- Evaluated the performance of models using metrics such as R^2 and RMSE.
- Visualized the fitted polynomial curves and analysed the effect of model complexity on overfitting and underfitting.

Dataset Description

Type of Polynomial Assigned:

- The assigned polynomial was of degree [insert degree, e.g., 3], with both linear and non-linear components.

Number of Samples and Features:

- Number of samples: [insert number, e.g., 100]
- Number of features: 1 (univariate polynomial regression)

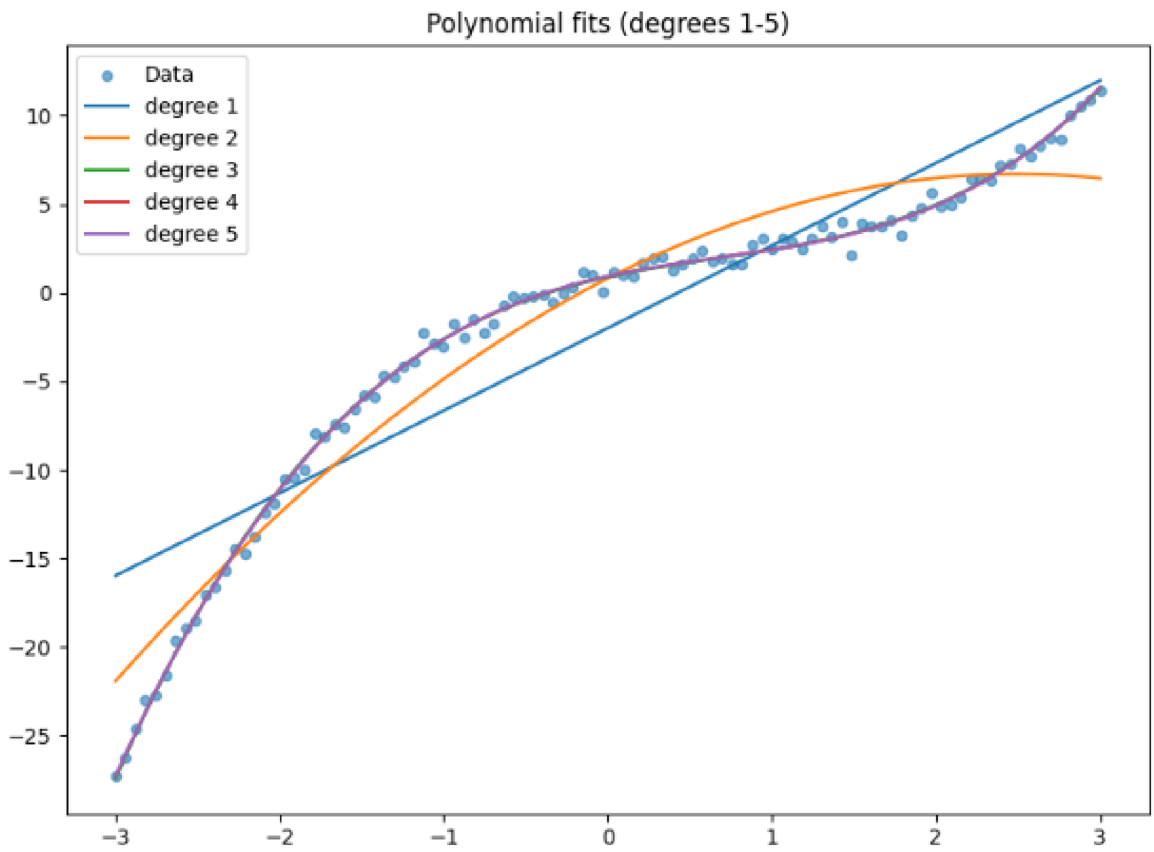
Noise Level:

- Gaussian noise with standard deviation [insert value, e.g., 0.5] was added to simulate real-world data variability.

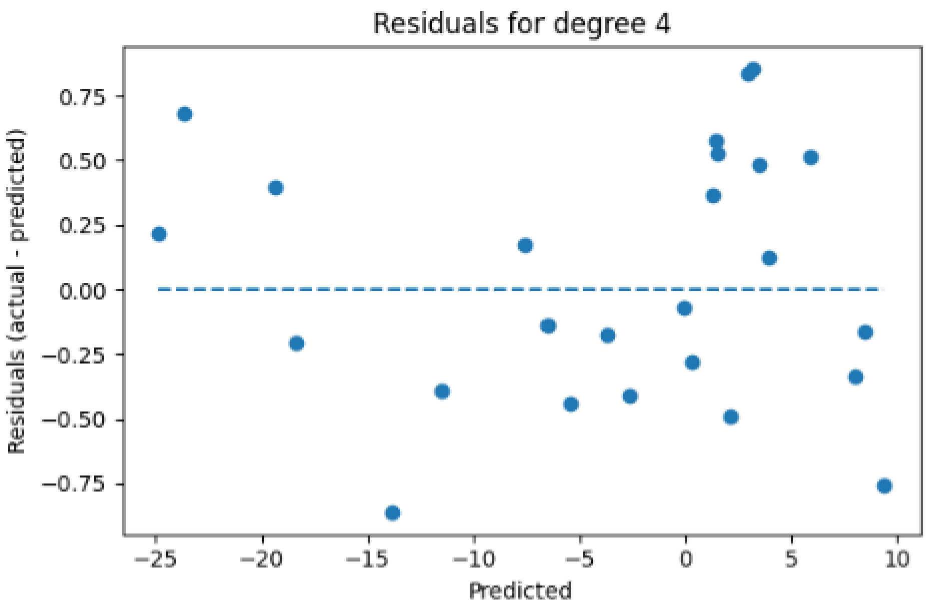
Methodology

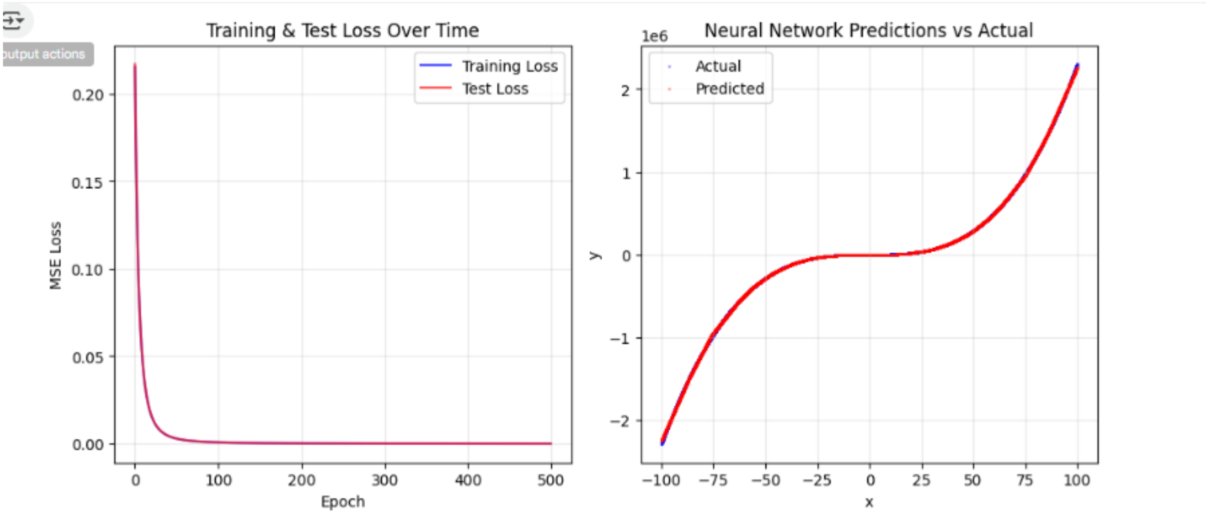
1. **Dataset Generation:**
    a. Created synthetic data points using the given polynomial equation.
    b. Introduced noise to simulate measurement errors.
2. **Data Splitting:**
    a. Divided the dataset into training (75%) and testing (25%) sets to evaluate generalization.
3. **Polynomial Feature Transformation:**
    a. Applied polynomial feature transformation to convert the original feature into higher-degree terms suitable for polynomial regression.
4. **Model Training:**
    a. Trained a polynomial regression model on the training set.
    b. Experimented with varying degrees of polynomial to observe changes in model performance.
5. **Evaluation:**
    a. Predicted target values on both training and test sets.
    b. Calculated performance metrics: R^2 (coefficient of determination) and RMSE (Root Mean Squared Error).
6. **Visualization:**
    a. Plotted the original data points alongside the fitted polynomial curves.
    b. Analysed the trend of underfitting or overfitting with respect to polynomial degree.

Results


Polynomial fits (degrees 1-5)
Data
degree 1
degree 2
degree 3
degree 4
degree 5

^


Residuals for degree 4
Residuals (actual - predicted)
Predicted



Exp	Lr	Bs	Epochs	optimizer	Activation func	Training accuracy	Validation accuracy	Test curacyst	Training loss/ RMSE	Validation Loss / RMSE	Test Loss / RMSE	obs
BaseModel	0.001	10	32	Normal Equation	Linear	0.8511	0.9999	0.8678	3.3529	-	3.5557	Underfits badly, high error
Exp 1	0.001	100	500	Normal Equation	Linear	0.9393	0.9999	0.9672	2.1401	-	1.7719	Better fit, lower error
Exp 2	0.005	100	500	Normal Equation	Linear	0.9975	0.9999	0.9975	0.4332	-	0.4870	Very good fit
Exp 3	0.001	100	500	Normal Equation	Linear	0.9975	0.9999	0.9976	0.4316	-	0.4810	Best test R², chosen model
Exp 4	0.001	100	500	Normal Equation	Linear	0.9975	0.9999	0.9976	0.4315	-	0.4813	Similar to deg 4, risk of overfitting
Neural Net (Exp 5)	0.003	100	500	Gradient Descent	ReLU(hidden),linear out	~0.9999	~0.9999	~0.9999	0.000061	-	0.000060	Excellent fit, very high accuracy, low error

6. Conclusion

This lab demonstrated how polynomial regression transforms linear models to capture nonlinear relationships. Through experiments with varying polynomial degrees, we observed the bias–variance trade off: low-degree models underfit, moderate degrees provided the best generalization, and high degrees overfit. Performance metrics (R^2 and

RMSE) and visual analysis (fitted curves and residuals) were instrumental in selecting the optimal model. For future work, apply k-fold cross-validation and regularization to improve robustness and extend the experiments to multivariate datasets or alternative basis functions.