What's the difference between polynomial/linear regression and logistic regression? Explain when you would use each one and how their outputs are different.

- Linear Regression predicts a continuous value (e.g., price, temperature). It assumes a straight-line relationship between features and the target.
- Polynomial Regression is a type of linear regression, but it adds polynomial terms (e.g., x^2, x^3) to model nonlinear relationships.
- Logistic Regression predicts the probability of belonging to a class (e.g., fraud vs. non-fraud) and outputs values between 0 and 1, which can be turned into class labels.

When we would use

Linear Regression: use when the target is continuous and the relationship is roughly straight-line. Example: predicting house prices from size.

Polynomial Regression: use when the target is continuous but the relationship is curved/nonlinear. Example: modeling population growth with time.

Logistic Regression: use when the target is categorical (yes/no, 0/1). Example: fraud detection, disease diagnosis, spam filtering.

output differences,

Linear / Polynomial Regression: output is a number anywhere from $-\infty$ to $+\infty$. For classification, we should set a threshold (e.g., ≥ 0.5 = fraud).

Logistic Regression: output is a probability (0–1) using the sigmoid function. No extra thresholding math is needed beyond the default 0.5.

Fraud dataset example results:

Model	Precision (Fraud)	Recall (Fraud)	F1-score (Fraud)	AUC	Notes
Logistic Regression	0.95	0.80	0.87	0.9829	Best balance of precision and recall for fraud detection.
Linear Regression	0.99	0.56	0.72	0.9763	Very high precision, but misses many fraud cases (low recall).
Polynomial Regression	0.95	0.82	0.88	0.9653	Slightly better recall than logistic regression, but AUC is lower.

I tested and compared three models: logistic regression, linear regression, and polynomial regression on the fraud dataset. Logistic regression worked best for this binary classification task, with an AUC of 0.9829 and a good balance of precision (0.95) and recall (0.80). Linear regression had a similar AUC (0.9763) and very high precision (0.99), but recall was much lower (0.56), meaning it missed many fraud cases. Polynomial regression (degree 2) gave slightly better recall

(0.82) than logistic regression, but its AUC was lower (0.9653), making it less reliable overall. These results showed that logistic regression is the most suitable choice for fraud detection in this case.