Calories Prediction Model

1. Problem Description

I built a model to **predict the number of calories burnt during an exercise session** using user and workout information such as age, weight, duration, heart rate, and body temperature.

Why? Calorie tracking helps people plan workouts, manage weight, and personalize training programs.

1. Dataset

I used a **public fitness dataset from Kaggle** (calories exercise dataset).

Target: Calories.

- Gender (binary) – may affect metabolism
- Age – metabolism changes with age
- Height & Weight – body size affects calories burnt
- Duration – longer workouts usually burn more
- Heart_Rate – intensity signal
- Body_Temp – reflects exertion

I checked for missing values and applied **StandardScaler** to the numeric features to improve model stability.

3. Modeling Approach

- Started with **Multiple Linear Regression (degree = 1)** to see if a simple linear model was enough.
- Then created **Polynomial features (degrees 2–6)** for the numeric columns: Age, Height, Weight, Duration, Heart_Rate, Body_Temp.
- **Gender** was left as is (binary), so I did not expand it into polynomial terms.
- For each degree: fit on the training data, evaluate on the test data, and record **R^2** and **MSE**. I collected the metrics in a table to compare models.

4. Evaluation

Metrics used:

- **R^2 Score** (higher is better, max = 1)
- **Mean Squared Error (MSE)** (lower is better)

Result: Polynomial degree **5** gave the **highest Test R^2** and **lowest Test MSE**, so it was selected as the best model. (Degree 6 started to get slightly worse, so it didn't help.)

5. Reflection

Challenges: Knowing when a linear model was not enough, and choosing the right polynomial degree without overfitting.

What I learned: Polynomial regression can really improve accuracy, but the degree must be chosen carefully using metrics. Scaling helps when using polynomial features.

Next time: I'll add more evaluation metrics (like MAE), try to add categorical data as input features and maybe test other model types for comparison.

Degree	Train R^2	Test R^2	Train MSE	Test MSE
1	0.9671	0.9673	127.1152	132.0268
2	0.9948	0.9951	20.0034	19.9734
3	0.9964	0.9966	13.9729	13.5712
4	0.9969	0.997	12.0629	11.9305
5	**0.9972**	**0.9972**	**10.8911**	**11.4364**
6	0.9974	0.9967	10.104	13.3481