Research Paper: Gym Workout Habits Dataset

Title & Collection Method

Title: Gym Workout Habits and Weight Gain Dataset

Collection Method: I designed a survey questionnaire and asked 50 young adults (ages 18–35) at my local gym about their workout routines, diets, and weight progress. Each participant provided answers about workout duration, frequency, diet type, supplement usage, and whether they gained weight in the last 3 months.

---

Description of Features & Labels

Features (X):

1. Workout Hours per Session (numeric, 0–3 hrs)

2. Gym Visits per Week (numeric, 1–7)

3. Diet Type (categorical: Balanced, High-Protein, Fast Food, Mixed)

4. Supplement Usage (categorical: None, Protein Powder, Creatine, Multiple)

5. Sleep Hours per Day (numeric, 4–10 hrs)

Label (y):

Weight Gain (3 months) → Yes / No

This makes the problem a classification task (predicting whether the person gained weight or not).

---

Dataset Structure

Rows: 50 gym members (samples)

Columns: 6 (5 features + 1 label)

---

Sample Table (10 rows)

Workout Hours Visits/Week Diet Type Supplement Sleep Hours Weight Gain

1.5 4 High-Protein Protein Powder 7 Yes

0.5 2 Fast Food None 6 No

2.0 5 Balanced Creatine 8 Yes

1.0 3 Mixed None 5 No

2.5 6 High-Protein Multiple 9 Yes

1.0 4 Fast Food Protein Powder 6 No

2.0 5 Balanced None 7 Yes

0.8 2 Mixed None 6 No

2.2 6 High-Protein Creatine 8 Yes

1.3 3 Balanced Protein Powder 7 Yes

---

Quality Issues

Missing values: A few participants didn’t report sleep hours.

Categorical text: Diet Type and Supplements need encoding for ML.

Imbalance: More participants reported Weight Gain (30) than No Gain (20).

Inconsistencies: Some wrote “HP” instead of “High-Protein.”

---

Use Case

This dataset can be used to train a classification model to predict whether someone will gain weight based on their gym habits, sleep, and diet.

Possible algorithms: Logistic Regression, Decision Tree, Random Forest.

Fitness coaches can use insights to recommend workout & diet plans for clients.